THE ANNALS
of
MATHEMATICAL STATISTICS

THE ANNALS OF MATHEMATICAL STATISTICS IS AFFILIATED WITH THE AMERICAN STATISTICAL ASSOCIATION AND IS DEVOTED TO THE THEORY AND APPLICATION OF MATHEMATICAL STATISTICS

Volume VIII, 1937

PUBLISHED QUARTERLY
ANN ARBOR, MICHIGAN

Four Dollars per annum

Made in United States of America

Address: ANNALS OF MATHEMATICAL STATISTICS
Post Office Box 171, Ann Arbor, Michigan

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD.

CONTENTS OF VOLUME VIII

APPLICATIONS OF TWO OSCULATORY FORMULAS

By John L. Roberts

INTRODUCTION

The main purpose of this paper is to illustrate how Mr. Jenkins' osculatory formulas[1] (A) and (B) can be applied in a convenient manner. The first section of this paper will be little more than a summary of some of the formulas contained in the other three articles. The second section will contain the applications.

I. SOME MATHEMATICS OF THE FORMULAS

The Woolhouse notation will in this paper be used to stand for the differences of u_{x+n} which represents the given values of a function. The general formulas are

$$y_x = y_0 + x\Delta y_0 + \tfrac{1}{2}x(x-1)B + \tfrac{1}{6}x(x-1)(x-\tfrac{1}{2})C; \qquad (1)$$

and

$$y_x = u_0 + xa_1 + \tfrac{1}{2}x(x-1)B + \tfrac{1}{6}x(x-1)(x-\tfrac{1}{2})C. \qquad (2)$$

The special formulas belonging to (2) are

$$B = b - \tfrac{1}{6}d \text{ and } C = c_1 - \tfrac{1}{6}e_1\,, \qquad \text{(A)}$$

where b and d are defined by $b = \frac{1}{2}(b_0 + b_1)$ and by $d = \frac{1}{2}(d_0 + d_1)$; and

$$B = b \text{ and } C = 0. \qquad \text{(B)}$$

The special formulas belonging to (1) are

$$y_0 = u_0 + \tfrac{1}{4}b_0\,, \qquad B = b, \qquad \text{and } C = 0; \qquad \text{(C)}$$

and

$$y_0 = u_0 - \tfrac{1}{36}d_0\,, \qquad B = b - \tfrac{1}{6}d, \qquad \text{and } C = c_1 - \tfrac{1}{6}e_1\,. \qquad \text{(D)}$$

Formula (C) is equivalent to Mr. Jenkins' formula (A). Also (D) is equivalent to his formula (B).

[1] This paper presupposes a knowledge of three other articles. The first one by Mr. Wilmer A. Jenkins is entitled "Graduation Based on a Modification of Osculatory Interpolation," and is printed in the October 1927 issue of the Transactions of the Actuarial Society of America. The other two papers are mine. One of them is entitled "Some Practical Interpolation Formulas," and is printed in the September 1935 issue of these Annals. The other one entitled "A Family of Osculatory Formulas" is printed in the October 1935 issue of the Transactions.

II. APPLICATIONS OF (C) AND (D)

First, there is the problem of selecting suitable examples to which (C) and (D) can be applied. Secondly, we will then apply in a convenient manner the formulas to these examples.

The problem of selecting suitable examples will now be considered. "The non-reproducing characteristic of" formula (D) "raises the question of what will happen in the graduation of a series whose fourth differences are all positive, say. The answer is that the graduated series will lie everywhere below the observed points and that the observations will not be correctly represented by the interpolated series." On the other hand, if we select a series whose fourth differences change frequently in sign, (D) because of its non-reproducing characteristic has valuable smoothing possibilities. In like manner, (C) may be valuable when the second differences change frequently in sign. Mr. Jenkins gives at quinquennial ages rates of mortality which were graphically determined from the published American Men Ultimate Experience. Since the fourth differences of these rates change frequently in sign, we will apply (D) to a few of these rates. So far as I know no suitable actuarial examples have been found to which (C) can be applied. However, there is the possibility that (C) might be valuable in some sciences. Since I do not know of any suitable real example to which (C) can be applied, we will apply it to a trivial series whose second differences change frequently in sign.

We are now ready to apply in a convenient manner (C) and (D) to the examples selected in the preceding paragraph.

First, we will apply (C). I have in my other article applied (B) in a convenient manner. This method with little change can be applied to (C). If it is desired to apply (C) at either end of the table where values of u_x are not available for the calculation of the second differences, it can be assumed they vanish. It is convenient if S and S^2 represent respectively the major differences Δu_x and $\Delta^2 u_x$ in such a manner that they are arranged centrally in the working illustration. It is convenient if s and s^2 represent respectively the minor differences δy_x and $\delta^2 y_x$. The quantity y_0 can be computed by $y_0 = u_0 + \frac{1}{4}b_0$, and y_1 can be computed in like manner. Since we wish in the working illustration of (C) to interpolate four values between y_0 and y_1, the middle $s = \delta y_{.4} = .2\Delta y_0$, and $s^2 = .04B = .02(b_0 + b_1)$. We can by the use of the foregoing method apply (C) to suitable functions, whose given values can be represented by $f(r)$. Then, it follows from the definition of u_x that $f(r) = u_x$. It might prevent confusion if it is stated that x and r are related to each other in such a way that we always interpolate between y_0 and y_1. We shall now apply (C) to the case when $f(r)$ represents the trivial series shown at top of page 3.

Finally, we will apply (D). Mr. Henderson has applied (A) in a very convenient manner. His method with little change can be applied to (D). If it is desired to apply (D) at either end of the table where values of u_x are not available for the calculation of the differences required, it can be assumed that the fourth differences that can not be computed vanish, and the required

r	$f(r)$	S	S^2	y_x	s	s^3
0	1		0	1.0	.18	
1				1.18	.14	
2		1		1.32	.10	−.04
3				1.42	.06	
4				1.48	.02	
5	2		−2	1.5		
6				1.5		
7		−1		1.5	.00	.00
8				1.5		
9				1.5		
10	1		2	1.5	.02	
11				1.52	.06	
12		1		1.58	.10	.04
13				1.68	.14	
14				1.82	.18	
15	2		0	2.0		

differences can be filled in consistently with that assumption. It is convenient if S, S^2, and S^3 represent respectively the major differences Δu_x, $\Delta^2 u_x$, and $\Delta^3 u_x$ in such a manner that they are arranged centrally in the working illustration. It is convenient if s, s^2, and s^3 represent the minor differences so that by definition $s = s_x = \delta y_x$, $s^2 = s_x^2 = \delta^2 y_{x-.2}$, and $s^3 = \delta^3 y_x$. The first $s^2 = \delta^2 y_{-.2} = .04(b_0 - \frac{1}{6}d_0)$. The last $s^2 = \delta^2 y_{.8} = .04(b_1 - \frac{1}{6}d_1)$. The quantity y_0 can be computed by $y_0 = u_0 - \frac{1}{86}d_0$, and y_1 can be computed in like manner. The middle $s = \delta y_{.4} = .2\Delta y_0 - s^3$. We are now in position to apply (D) to the quinquennial rates of mortality.

Age	Rate	S	S^2	S^3	s^4
72	.07010				
		.03808			
77	.10818		.00861		
		.04669		.01799	
82	.15487		.02660		−.03745
		.07329		−.01946	
87	.22816		.00714		.14518
		.08043		.12572	
92	.30859		.13286		.00000
		.21329		.12572	
97	.52188		.25858		.00000

Age	y_x	δ	δ^2	δ^3
82	.15591	12612	.001314	
83	.168522	13527	915	
84	.182049	.014043	516	−.000399
85	.196092	14160	117	
86	.210252	13878	−.000282	
87	.22413	13460	−.000682	
88	.237590	13977	.000517	
89	.251567	.015693	1716	.001199
90	.267260	18608	2915	
91	.285868	22722	4114	
92	.30859	28006	.005314	
93	.336596	34326	6320	
94	.370922	.041652	7326	.001006
95	.412574	49984	8332	
96	.462558	59322	9338	
97	.52188		.010343	

SOME SIMPLE DEVELOPMENTS IN THE USE OF THE COEFFICIENT OF STABILITY

By C. H. Forsyth

Some time ago the writer proposed[1] a coefficient of stability C_s to be used to measure the stability of a statistical series, where that coefficient is defined by the relation

$$C_s = \frac{\sigma^2}{M} \tag{1}$$

where M denotes the arithmetic mean and σ^2 the square of the dispersion of the terms of the series. It was proposed to regard series as unstable (Lexian) for which the value of the coefficient exceeded unity, and stable otherwise. The only essential way in which such a procedure differs *in results* from the traditional method is that it includes as stable those series for which the value of the coefficient lies between unity and q the probability of failure of the event under investigation—series which would be classed as unstable according to the traditional method. Stable series—according to either standard—are found so rarely in practice and therefore so many series are accepted as fairly stable which come anywhere near meeting the requirements that replacing q by unity as the line of demarcation affects the classification of no known series but adds to the effectiveness of the avowed purpose and use of the proposed coefficient—to avoid the round-about work of computing values of probabilities. Another merit of the use of the coefficient is that it enables one to measure and therefore *compare* the stability of several series—a feature which we shall illustrate later.

In brief, such a coefficient provides a means of introducing the whole Lexian theory into Federal publications such as those on vital statistics, since a comparison of the values of the coefficient for, say different communities or countries, would be readily grasped by any reader, whereas the traditional method would prove too subtle and laborious, and allow no ready comparison of results.

For purpose of orientation let us illustrate the situation by analyzing a simple series both ways—the traditional way and by the use of the coefficient of stability. As an example, let us consider the death rates of white infants under one year of age for 1919 (considered on page 89 of the Handbook) for those states whose frequencies of births are comparable or which vary little from

[1] Journal of the American Statistical Association, June, 1932.

their average of 47,830—where the number of deaths for each state has been adjusted to this average as a base.

	Adjusted Deaths X	$X - 3659$	$(X - 3659)^2$
Cal.	3350	−309	95481
Conn.	4700	1041	1083681
Ind.	3732	73	5329
Kan.	3253	−406	164836
Ky.	3686	27	729
Minn.	3159	−500	250000
N. Car.	3541	−118	13924
Va.	3732	73	5329
Wis.	3780	121	14641
	9)32933	1335–1333	)1633950
	$M = 3659$		$181550 = \sigma^2$
			$\sigma = 426$

The traditional method would be:

The mean $M = np = 3659$ where $n = 47{,}830$.

$$\text{Hence } p = \frac{3659}{47830} \text{ and } q = \frac{44171}{47830}$$

$$\text{and } \sigma_B^2 = npq = 3659\left(\frac{44171}{47830}\right) = 3378$$

$$\text{whence } \sigma_B = 58.15$$

which is the value of the dispersion we should expect if the basic probability were constant throughout. But the value of the dispersion proves to be $\sigma = \sqrt{181550} = 426$, and the comparison of the values shows that the basic probability to be very variable and therefore the series to be very unstable or Lexian.

The computation of the value of the coefficient of stability is much more simple and direct

$$C_s = \frac{\sigma^2}{M} = \frac{181550}{3659} = 49.6$$

whose excess over unity also clearly indicates the instability of the series.

Since proposing the coefficient of stability the writer has been impressed by the overwhelming proportion of existing series (such as birth rates, various kinds of death rates, etc.) which employ arbitrary bases (such as "per thousand," "per ten thousand," etc.) usually without mention of the actual base. It is obvious, of course, that such rates, or occurrences per arbitrary base, say b, can first be adjusted to give occurrences per actual base, say B (assuming that

base B* can be determined) but the work can evidently be performed much easier. For, since the original series (per arbitrary base b) $X_1, X_2, \cdots X_N$ would become, on adjustment, $\frac{B}{b}X_1, \frac{B}{b}X_2, \cdots \frac{B}{b}X_N$, the mean would become $\frac{B}{b}M$ and the square of the dispersion $\left(\frac{B}{b}\sigma\right)^2$, whence the formula for the coefficient of stability would become

$$C_s = \frac{\sigma^2}{M} \cdot \frac{B}{b} \tag{2}$$

As an example, let us consider the general death rates, per 10,000, of New Zealand for the years 1921–30.

	X	$X - 86$	$(X - 86)^2$
1921	87	1	1
1922	88	2	4
1923	90	4	16
1924	83	−3	9
1925	83	−3	9
1926	87	1	1
1927	85	−1	1
1928	85	−1	1
1929	88	2	4
1930	86	0	0
	10)862	10−8	)46
	M = 86.2		4.6

This example illustrates the danger of using the coefficient of stability unless the series consists of actual occurrences or unless the actual base is given due consideration. Without due consideration of the actual base (here the population of New Zealand) one might easily fall into the error of regarding the value of the coefficient of stability as 4.6/86.2 and, therefore, the series as very stable. But the population of New Zealand is about a million and a half and, therefore the true value of the coefficient of stability is

$$C_s = \frac{4.6}{86.2} \frac{1,500,000}{10,000} = 8.0$$

* Strictly speaking, this actual base B should be constant throughout the series; otherwise the successive numbers of occurrences—the terms of the series—would not be comparable. Where, however, the base B varies little from term to term—as usually happens even in the best of series, such as a series of some kind of rates of the same community over a short interval—the variation can be ignored, in which case base B (to which the terms of the series are adjusted) usually means the arithmetic mean of the different bases. In the first treated above, the investigation was limited to certain states in an effort to comply with the rule just mentioned but the example is a poor one since the variations are still dangerously too large. The situation is saved by the conclusive results.

which shows the series to be unstable. However, before we condemn New Zealand's death rates too severely, let us compare her record with those of other important countries, including our own, for the same period.

General Death Rates (per 10,000)

	M	C_s
New Zealand	86.2	8
Australia	94.3	90
Sweden	120.4	96
Scotland	137.3	139
Austria	151.1	536
United States	118.0	830
England-Wales	121.3	1117
France	170.3	1129
Spain	193.7	2190
Italy	163.5	2760
Germany	125.4	6040
Japan	206.4	6800

These results show how extremely unstable most series of general death rates are and that the series for New Zealand, while unstable according to our strict criterion, enjoys quite an enviable position practically in a class by itself. Parenthetically, these results also illustrate fairly well the triviality, with respect to results, of replacing q by unity as the critical value of the coefficient of stability, discussed at the beginning of this article.

The values of the coefficient listed above would, of course, be reduced somewhat in most cases if the trend of the series were first eliminated but the writer has gone though all this work and found it not worth while—that is, the series would still remain markedly unstable.

Another development proves useful when, as frequently happens, the actual base B is unknown to a degree of accuracy desirable for use in formula (2).

From the inequality $\frac{\sigma^2}{M}\frac{B}{b} \leqq 1$

we obtain

$$B \leqq \frac{Mb}{\sigma^2} \tag{3}$$

which is to be used to show how small an actual base should be for the given series to be stable. As an example, let us consider the maternal mortality, per 10,000 live births, in the so-called expanding registration area of the United States.

Maternal Deaths in the United States (per 10,000 live births) (Expanding Registration Area)

	X	$X - 66$	$(X - 66)^2$
1923	67	1	1
1924	66	0	0
1925	65	−1	1
1926	66	0	0
1927	65	−1	1
1928	69	3	9
1929	70	4	16
1930	67	1	1
1931	66	0	0
1932	64	−2	4
	10)665	9−4	)33
	66.5		3.3

Hence, by formula (3), $B \leqq \frac{66.5}{3.3}$ (10,000) or about 200,000. The number of live births varies so greatly that we should probably find it impossible to agree upon a satisfactory number[2] to use as an actual base for such an "expanding area" but we should all agree that it would be so much greater than 200,000 that the instability of the series would be unquestioned.

One must be careful in comparing the results of two or more investigations like the one just conducted. For example, the analogous result for Canada for the same period yields $B \leqq 113{,}000$ and we might conclude, too hastily, that the United States series is more stable (or less unstable) whereas any knowledge whatever of the numbers of live births of the two countries would show that Canada comes much closer to fulfilling her requirement than the United States and that the palm must go to Canada. For one thing, Canada has about the population of New York city and New York city has about 100,000 live births annually. In any case, *close* decisions in matters of this kind would be difficult without sufficient information in regard to actual bases.

There is still another situation which is interesting but of much less importance because of the rarity of its occurrence. It will be recalled that the coefficient of stability was devised mainly to avoid the use and computation of probabilities and that the only difference between the results by the traditional method and by the use of the coefficient of stability lies in the trivial replacement of the critical value q by unity. In the traditional method of analysis, but by comparing the value of the coefficient of stability with q, the coefficient is evidently always, strictly speaking, a function of the actual base B. In other words, there is no statistical series, however stable it may seem—except

[2] It was in the neighborhood of two million in 1932.

for the trivial case when all the terms of the series are exactly the same—but what would be unstable if the base were small enough. It is possible to formulate the limit once for all below which the given (otherwise seemingly stable) series would prove unstable.

If, in the relation $\sigma^2 \leqq npq$ (for stability) we replace p by M/n, q by $1 - M/n$ and then n by B, we obtain

$$\sigma^2 \leqq M - \frac{M^2}{B} \text{ or } \frac{M^2}{B} \leqq M - \sigma^2$$

whence, finally

$$B \geqq \frac{M^2}{M - \sigma^2} \tag{4}$$

where the transference of the term $M - \sigma^2$ from one side to the other should cause no apprehension since, by hypothesis, $\sigma^2 < M$ and $M - \sigma^2$ is therefore always positive. We propose to employ formula (4) in those rare cases where the value of the coefficient of stability of actual occurrences—but without reference to an actual base—is less than unity—that is, where the given series proves to be stable according to the method proposed by the writer—and determine the upper limit of the values of the base B for which the series would be unstable according to the traditional method of analysis. As an illustration, let us consider the familiar series of annual football fatalities in this country for the period 1906–1930* (omitting the years when no records were kept).

Football Fatalities

1906	11	1917	12
1907	11	1921	12
1908	13	1923	18
1909	12	1925	20
1911	11	1926	9
1912	13	1927	17
1913	5	1928	18
1914	13	1929	12
1915	15	1930	13

It is easily verified that $C_s = \dfrac{11.942}{13.055}$ which is clearly less than unity; whence the series clearly seems stable. Applying formula (4)

$$B \geqq \frac{13.055^2}{13.055 - 11.942} \quad \text{or } 153$$

which shows that the given series is stable as long as the total number of football players exceeds the number 153. A recent news item quoted an estimate of the number players participating in games of four hundred colleges as about

13,000 and over 600,000 including high schools and all. We can then definitely say that the series just considered is stable. Such a conclusion has no bearing, of course, upon what might happen if other terms were added to the series. It happens that adding the records for the next five years—1931(33), 1932(32), 1933(27), 1934(25), 1935(30)—would change the whole series to an unstable one with $C_s = 56.9/16.6 = 3.4$; but, obviously, the additional records belong to a new regime of collection.

INTERNAL AND EXTERNAL MEANS ARISING FROM THE SCALING OF FREQUENCY FUNCTIONS

By Edward L. Dodd

The scaling[1] of frequency functions has been discussed from the standpoint of maximum likelihood. But the likelihood criterion to be satisfied sometimes leads to a minimum likelihood; and sometimes to neither a maximum nor a minimum. Scaling will be studied in this paper with reference to the likelihood actually secured, and also with reference to the character of means obtained, whether internal or external.

SECTION 1. INTRODUCTION

It is well known that a scale obtained in a curve-fitting process is sometimes a mean. Thus, with the normal function

$$\frac{1}{a\sqrt{2\pi}}\, e^{-(x/a)^2/2}, \tag{1}$$

if the scale a is to be obtained from measurements, $x_1, x_2, \cdots, x_n$, we commonly accept the value

$$a = \left\{\frac{1}{n}\sum x_i^2\right\}^{1/2}; \tag{2}$$

that is, the root-mean square of the measurements. Here, the positive value of a is naturally taken. It is called the standard deviation, and thought of as an appropriate new unit of measure.

But even with the x's all negative, and the a taken positive, O. Chisini[2] considered it proper to regard a as a mean of the x's, albeit an *external* mean. From Chisini's viewpoint, this a whether regarded as positive or negative is primarily a solution of

$$x_1^2 + x_2^2 + \cdots + x_n^2 = a^2 + a^2 + \cdots + a^2. \tag{3}$$

In this sum of squares, the single number a may be *substituted* for each of the x's. Perhaps this kind of mean should be called a *substitutive* mean to distinguish it from the means of general analysis which are always internal.

[1] Fisher, R. A., "On the mathematical foundation of theoretical statistics," Philosophical Transactions of the Royal Society of London, Series A, Vol. 222, 309–368, (1921). See p. 338.

[2] Chisini, O., "Sul concetto di media," Periodico di matematico, Series 4, Vol. 9, 106–116, (1929).

The normal function is a particular case of a more general function:

$$\text{Constant}\cdot a^{-1}\, e^{\phi(t)}, \qquad \phi(t) = -t^p/p, \qquad t = x/a. \tag{4}$$

The likelihood method to find the scale a for this function leads to power means, including the arithmetic mean, the root-mean-square, root-mean-cube, etc., for $p = 1, 2, 3$, etc.

The word *scale* will be used only for a positive number,—which then may be regarded as a unit of measurement.

For measurements, $x_1, x_2, \cdots, x_n$ Chisini regarded M as a mean, relative to a function G, provided

$$G(x_1, x_2, \cdots, x_n) = G(M, M, \cdots, M). \tag{5}$$

If a solution of this equation is

$$M = F(x_1, x_2, \cdots, x_n), \tag{6}$$

and c is a possible value for the x's, it follows at once that

$$F(c, c, \cdots, c) = c, \tag{7}$$

or at least one value of this F is c. Conversely, if (7) is satisfied, it is but a change of notation to replace c in (7) by M, and to combine this with (6) to obtain

$$F(x_1, x_2, \cdots, x_n) = F(M, M, \cdots, M). \tag{8}$$

Hence, this F which in (6) gives explicit form to the implicit M found in (5) may also be thought of as a mean-forming function, such as G in (5). Briefly, F is a particular G. Thus $F(x_1, x_2, \cdots x_n)$ is a mean of $x_1, x_2, \cdots, x_n$, if F is so constructed that (7) is satisfied when the arguments are all equal.

Inasmuch as a frequency function $f(t)$ is non-negative, $\log_e f(t)$ is real,—say $\phi(t)$ plus constant. Following R. A. Fisher, it will be convenient to write

$$f(t) = Ca^{-1}\, e^{\phi(t)}, \qquad C = \text{Constant} \tag{9}$$

With location m already determined, the x's will be thought of as measured from m. And we set

$$t = x/a, \qquad t_i = x_i/a, \qquad i = 1, 2, \cdots, n. \tag{10}$$

The "productive" probability—to yield $x_1, x_2, \cdots, x_n$—is then

$$L = \Pi f(t_i) = C^n\, a^{-n}\, e^{\Sigma\phi(t_i)}. \tag{11}$$

This is proportional[3] to the "likelihood" of a. Also—it may be noted in passing—the productive probability is also proportional to the *a posteriori* probability, if a constant *a priori* probability is postulated. The likelihood will here be taken as $\Pi f(t_i)$ itself; and it will be designated by L,—in Fisher's

[3] Loc. Cit., Fisher, p. 310.

notation, $L = \log \Pi$. Of course, Π and $\log \Pi$ take maximum values simultaneously, if at all. From (11) it follows[4] that

$$-a \cdot \partial \log L/\partial a = n + \Sigma t_i\phi'(t_i) = \Sigma\{t_i\phi'(t_i) + 1\}. \tag{12}$$

The equation

$$\Sigma t_i\phi'(t_i) + n = 0 \qquad (i = 1, 2, \cdots, n) \tag{13}$$

will be called the likelihood condition, whether this leads to maximum likelihood, to minimum likelihood, or to neither. A second differention[5] leads to

$$a^2 \cdot \partial^2 \log L/\partial a^2 = \Sigma t_i^2\phi''(t_i) - n = \Sigma\{t_i^2\phi''(t_i) - 1\}. \tag{14}$$

When negative, this indicates a maximum likelihood; when positive, a minimum likelihood for the a obtained from (13).

Preparatory to the theorems of the next section, just one more matter will be discussed. The unit for t is arbitrary; and it may be convenient to write, with $k \neq 0$,

$$\phi(t) = \phi(ku) = \Phi(u), \qquad t = ku. \tag{15}$$

Then

$$t\phi'(t) = u\Phi'(u). \tag{16}$$

Suppose, now, that a positive constant k can be found such that $k\phi'(k) = -1$. Then, with $t = ku$, as postulated,

$$1 \cdot \Phi'(1) = k\phi'(k) = -1. \tag{17}$$

Thus $\Phi'(1) = -1$,—or as it will now be written $\phi'(1) = -1$,—is no more restrictive than the condition that some positive k exists such that $k\phi'(k) = -1$.

SECTION 2. GENERAL THEOREMS CONCERNING THE SCALE AS A MEAN

Theorem I

Given the frequency function

$$f(t) = Ca^{-1} e^{\phi(t)}, \qquad t = x/a, \qquad t_i = x_i/a, \qquad C = \text{Constant}. \tag{18}$$

And suppose that

$$\phi'(1) = -1. \tag{19}$$

Suppose, also, that for given $x_1, x_2, \cdots, x_n$, the likelihood condition (13), now written

$$\textstyle\sum_1^n (x_i/a)\phi'(x_i/a) + n = 0, \tag{20}$$

[4] Loc. Cit., Fisher, p. 338.
[5] Loc. Cit., Fisher, p. 339.

has a positive solution,

$$(21) \qquad a = F(x_1, x_2, \cdots, x_n).$$

Then this a, the scale, is a mean.

Proof. With each $x_i = 0$, (20) cannot be satisfied.

But if, with $c \neq 0$, we take each $x_i = c$, and at the same time set $a = c$, then, by (19), $\Sigma = -n$; and thus (20) which gives a implicitly is satisfied. The explicit a in (21) is therefore such a function F that (7) is satisfied. Hence, the scale a is a mean.

THEOREM II

Given the frequency function

$$(18) \qquad f(t) = Ca^{-1} e^{\phi(t)}, \qquad t = x/a, \qquad t_i = x_i/a, \qquad C = \text{Constant}.$$

Suppose that

$$(19) \qquad \phi'(1) = -1,$$

and that

$$(22) \qquad |t\phi'(t)| < 1 \qquad \text{if} \qquad |t| < 1.$$

Moreover, suppose that the likelihood condition (20) for measurements $x_1, x_2, \cdots, x_n$, has a positive solution a. Then

$$(23) \qquad a \leqq \text{Maximum } |x_i|.$$

Or, suppose that, in place of (22), we have

$$(24) \qquad |t\phi'(t)| > 1 \qquad \text{if} \qquad |t| > 1;$$

and that $t\phi'(t)$ keeps the same sign, if $|t| > 1$. Then

$$(25) \qquad \text{Minimum } |x_i| \leqq a.$$

Proof. Suppose, if possible, that $a > \text{Max } |x_i|$. Then each $|x_i/a| < 1$, and by (22), $|(x_i/a)\phi'(x_i/a)| < 1$. Then (20) is not satisfied, since $|\Sigma| < n$. Thus the hypothesis is contradicted.

Now (25) is satisfied at once if any $x_i = 0$. But suppose, on the other hand, that $\text{Min } |x_i| > 0$; and, if possible, that $a < \text{Min } |x_i|$. Then, by (24) et seq., since $|x_i/a| > 1$, it follows that $|\Sigma| > n$. And thus (20) is again contradicted.

THEOREM III

Given the frequency function

$$(18) \qquad f(t) = Ca^{-1} e^{\phi(t)}, \qquad t = x/a, \qquad t_i = x_i/a, \qquad C = \text{Constant};$$

and set $\psi(t) = t\phi'(t) + 1$. Suppose that

$$(26) \qquad \lim_{t \to 0} \psi(t) = \alpha, \qquad \lim_{|t| \to \infty} \psi(t) = \beta, \qquad \alpha\beta < 0.$$

And suppose that $\psi(t)$ is continuous when $t \neq 0$.

Then, for any set of real numbers, $x_1, x_2, \cdots, x_n$, of which none is zero, there exists a positive number a, as scale, such that the likelihood condition

$$\sum_1^n (x_i/a)\,\phi'(x_i/a) + n = 0 \tag{20}$$

is satisfied.

The conclusion is also valid, if in place of the limit β, there is postulated

$$\lim_{t \to -b+0} \psi(t) = -\alpha \mid \infty \mid = \lim_{t \to c-0} \psi(t), \tag{27}$$

where $b > 0$, $c > 0$, and $\psi(t)$ is continuous for $-b < t < 0$ and for $0 < t < c$. That is, the new limits are to be infinite with sign opposite to that of α.

Proof. The limits for $t \to 0$ and for $\mid t \mid \to \infty$ are the same as the limits for $a \to \infty$ and $a \to 0+$,—noting that $t = x/a$, $x \neq 0$. Thus $\Sigma\psi(t_i)$ changes sign as a goes from $0+$ to ∞. Hence, since $\psi(t)$ is continuous, (20) is satisfied for some positive a.

For the proof of the second part of the theorem, suppose that $x_n > 0$ and that x_n is the greatest x_i. Then with $a > x_n/c$, but approaching x_n/c, $\psi(x_n/a)$ becomes infinite with sign opposite to that of α. Furthermore, in $\Sigma\psi(x_i/a)$, the positive x's $< x_n$ have a negligible effect; and thus $\lim \Sigma\psi(x_i/a)$, as $a \to (x_n/c) + 0$, is infinite with sign opposite to that of α, when this sum Σ is taken for the positive x's. Likewise, if $x_1 < 0$, and is the least x_i, $\lim \Sigma\psi(x_i/a)$, as $a \to (-x_1/b) + 0$, is infinite with sign opposite to that of α, when this sum is taken for the negative x's. If, now, the measurements happen to be all positive, we think of a as approaching $x_n/c + 0$; and the continuity condition leads to an a which makes $\Sigma\psi(x_i/a) = 0$. Likewise, if the measurements happen to be all negative, we use $-x_1/b + 0$. If both positive and negative x's appear, we use the greater of the two ratios $-x_1/b$ and x_n/c.

SECTION 3. SOME FAIRLY REGULAR FREQUENCY FUNCTIONS

To illustrate the foregoing theorems in a somewhat general manner, consider the measurements, $x_1, x_2, \cdots, x_n$, and with $t = x/a$, $t_i = x_i/a$, set up the function:

$$f(t) = Ca^{-1} \mid kt \mid^p (1 + k^2t^2)^{-q}\, e^{-r \mid kt \mid^s}, \tag{28}$$

where, as before, C is a suitably chosen constant.

Suppose also that

$$p > -1, \qquad q \geqq 0, \qquad r \geqq 0, \qquad s \geqq 0; \tag{29}$$

and that either

$$r > 0,\ s > 0 \qquad \text{or} \qquad r = 0,\ 2q > p + 1. \tag{30}$$

Then with $\phi(t) = \log f(t)$, it follows that, when $t \neq 0$,

$$t\phi'(t) + 1 = (p + 1) - rsk^s \mid t \mid^s - 2qk^2t^2(1 + k^2t^2)^{-1}. \tag{31}$$

Now the condition $1 \cdot \phi'(1) = -1$ would be satisfied if $\Psi(k) = 0$, where

$$(32) \qquad \Psi(k) = rsk^{s+2} + rsk^{s} + (2q - p - 1)k^2 - (p + 1).$$

But, under the conditions (29) and (30) $\Psi(0) < 0$, and $\Psi(\infty) > 0$. Hence, there is a positive k for which $\Psi(k) = 0$. Then if k be assigned this value, (19) is satisfied; and by Theorem I, any scale a that the likelihood condition (20) may lead to is a mean. But, by Theorem III a scale a will actually exist —indeed, for any positive k that may be used in (29); since the limit of $t\phi'(t) + 1$ is positive as $t \to 0$, and is negative as $|t| \to \infty$.

Moreover, if in (29), the further condition $-1 < p \leqq 0$ is introduced, (22) is satisfied. And, thus, $a \leqq$ Maximum $|x_i|$. Also, $|t\phi'(t)|$ increases with $|t|$. Hence, by (24) et seq., Minimum $|x_i| \leqq a$.

If in (28), we set $q = 0$, $s = 1$, $r > 0$, and confine our attention to positive x and t, there is obtained the Pearson Type III. Reference to (32) shows that $\Psi(k) = 0$ if $k = (p + 1)/r$. With this substitution,

$$(33) \qquad f(t) = C^1 a^{-1} t^p e^{-(p+1)t}, \qquad C' = \text{Constant}.$$

Since $\phi'(1) = -1$, any solution of the likelihood condition is a mean. Here, with $t > 0$, $t\phi'(t) = p - (p + 1)t$, and $t^2\phi''(t) - 1 = -(p + 1)$. From (14) we see that, with $p + 1 > 0$, any mean obtained corresponds to maximum likelihood and the single maximum found is actually the largest value. Moreover, with the measurements, $x_1, x_2, \cdots, x_n$, all positive, a scale a will exist,—as noted in the general case (28).

In passing, it may be noted that Type III appears[6] rather naturally in a form giving $\phi'(1) = -1$ at once, without any transformation. Here, then, a scale is a mean.

Given the Pearson Type I in the form

$$(34) \quad f(t) = Ca^{-1}(b + kt)^p(c - kt)^q, \quad t = x/a, \quad b > 0, \quad c > 0, \quad |pq| > 0.$$

If $p + q + 1 > 0$, it is possible to find a positive k so that with $\phi = \log f$, $\phi'(1) = -1$. In this case, any scale found by the likelihood condition is a mean. With k thus chosen, $f(t)$ has essentially the same form as it would have if $k = 1$. Hence for convenience, let us simply set $k = 1$ in the above equation. Then for $-b < t < c$,

$$\psi(t) = t\phi'(t) + 1 = 1 + pt(b + t)^{-1} - qt(c - t)^{-1}.$$

Suppose now that $p > 0$ and $q > 0$. Then Theorem III may be applied; since $\lim \psi(t) = 1$, as $t \to 0$; but $\lim \psi(t) \to -\infty$, as $t \to -b + 0$, or as $t \to c - 0$.

[6] Carver, H. C., Handbook of Mathematical Statistics, Chap. VII, see p. 105, Line 4, noting that $\phi' = y'/y$.

Hence a scale a satisfying the likelihood condition exists. Moreover, the likelihood is at a maximum; since, with $-b < t < c$,

$$t^2\phi''(t) - 1 = -pt^2(b + t)^{-2} - qt^2(c - t)^{-2} - 1 < 0.$$

This maximum is also the largest value for all values of a.

If the Pearson Type IV is given in the form

$$f(t) = Ca^{-1}(1 + k^2t^2)^{-p}\, e^{q \arctan kt}, \qquad t = x/a \tag{35}$$

then if $p > 1/2$, it is possible to find a positive k which will make $\phi'(1) = -1$. In this case, any scale a is a mean. Moreover—for any $k \neq 0$—the limit of $t\phi'(t) + 1$ is 1 for $t \to 0$ and is $1 - 2p$ for $t \to \infty$. Hence, by Theorem III, if $p > 1/2$, as above, then a scale a exists satisfying the likelihood condition (20).

SECTION 4. FREQUENCY FUNCTIONS WITH CERTAIN PECULIARITIES

The theorems of section 2 give sufficient conditions, which in some cases may not be necessary. Nevertheless, by violating certain hypotheses, particular functions may be set up which exhibit various peculiarities.

For the Pearson Types, the differential equation is

$$\frac{y'(t)}{y(t)} = \phi'(t) = \frac{a_0 + a_1 t}{b_0 + b_1^t + b_2 t^2}, \qquad t = x/a. \tag{36}$$

The determination of a positive scale a by the Fisher likelihood process is impossible here, in case $a_0 = 0$, $a_1 > 0$, $b_0 + b_1 t + b_2 t^2 > 0$. For in this case $t\phi'(t) \geqq 0$; and thus (20) cannot be satisfied. The U-shaped Type II curves are in this class. Likewise, if $a_0 \neq 0$, $a_1 = 0$, and $b_0 + b_1 t + b_2 t^2 > 0$,—for example, with $b_2 > 0$, $b_1^2 < 4b_0 b_2$,—and the measurements all happen to have the same sign as a_0, such scaling is impossible.

For the purpose of constructing peculiar functions we may take $c > 0$ and require that the measurements x_i be either $-c$ or c—with at least one $-c$ and at least one c—and that $\phi(t)$ be an even function. Then $\phi(-c) = \phi(c)$ and (11) becomes

$$L = [Ca^{-1}\, e^{\phi(c/a)}]^n. \tag{37}$$

The likelihood condition (13) reduces to

$$0 = \psi(t) = t\phi'(t) + 1 = (c/a)\phi'(c/a) + 1, \tag{38}$$

with the right member an even function of c/a. And from (14), a maximum likelihood is indicated when

$$(c/a)^2\phi''(c/a) - 1 < 0, \tag{39}$$

with the left member likewise an even function. A minimum likelihood is indicated if the left member is positive.

Let us apply this to the case where

(40) $\phi(t) = (-2/3) \log (1 - 3 \mid t \mid); \quad t\phi'(t) = 2 \mid t \mid (1 - 3 \mid t \mid)^{-1}.$

The likelihood condition (38) is satisfied only when $t = \pm 1$. Also $\phi'(1) = -1$. Thus the only means are the internal means $\pm c$; and the only scale conformable to (38) is $a = c$. But this has minimum likelihood; since $1 \cdot \phi''(1) - 1 = \frac{1}{2} > 0$. For positive t, this function (40) is a Pearson Type.

Consider next a function of the form (28),—with $p = -1.25$, $q = -0.5$, however,—for which (31) becomes

(41) $t\phi'(t) + 1 = -1/4 - t^2/4 + t^2/(1 + t^2) = -(1 - t^2)^2/4(1 + t^2).$

whence $\phi'(1) = -1, \phi''(1) = +1, \phi'''(1) = -3$. Here the likelihood condition (38) has but a single absolute solution $\mid t \mid = 1$, leading to the single scale $a = c$, and to the two internal means, $\pm c$. But, in this case $1 \cdot \phi''(1) - 1 = 0$, so that $\partial^2 \log L/\partial a^2 = 0$. Moreover, for $t = 1$, $\partial^3 \log L/\partial a^3 = a^{-3} \neq 0$. Thus, the only scale obtained by the likelihood method (38)—viz., $a = c$— has a likelihood which is neither at a maximum nor at a minimum.

Another anomalous function is that given by

(42) $$\phi(t) = t^4 - 2.5t^2, \qquad t = \pm c/a.$$

The likelihood condition (38) leads to

$$\psi(t) = (1 - t^2)(1 - 4t^2) = 0.$$

The only solutions are $t = \pm 1$, giving internal means $\pm c$; and $t = \pm 1/2$, giving external means $\pm 2c$. And from (39) et seq., it can be shown that the internal mean and scale, $a = c$ has minimum likelihood, while the external mean and scale, $a = 2c$, has maximum likelihood.

But it will be noted that a maximum value for a vicinity does not always signify a largest value for the entire possible range. Indeed, for the function (42), $a = 2c$ has maximum likelihood without having the largest likelihood. To avoid such an anomaly, a necessary condition is that as $\mid t \mid \to \infty$, $\psi(t) \to -\infty$; as seen by taking the logarithm of L in (37), noting that as $a \to 0$, $(-\log a) \to +\infty$.

Finally avoiding the anomaly just mentioned, let us set up a frequency function, using the $\psi(t)$ in (38), and writing

$$\psi(t) = 1 + t\phi'(t) = (1 - 2t^2)(1 - t^2)(1 - 0.9t^2).$$

From this it follows readily that

(43) $$\phi(t) = K - 1.95t^2 + 1.175t^4 - 0.3t^6, \qquad K = \text{Constant}.$$

This, with $t_i = \pm c/a$, leads to an internal mean or scale $a = c$ with minimum likelihood, a nearby scale $a = c\sqrt{0.9}$ with maximum likelihood—differing indeed only slightly from the minimum just mentioned—and another scale $a = c\sqrt{2}$ having maximum likelihood, and this likelihood is indeed greater

than that for any other positive value of a. The external mean $a = c\sqrt{2}$ in this case has the largest likelihood. This may be checked by the use of the logarithm of L as it appears in (37), in which the important part is $\phi(c/a) - \log a$.

In passing it may be noted that if $\psi(t)$ has the form $\psi(t) = (1 - t)H(t)$, with $H(1) \neq \infty$, and $t_i = x_i/a$; then any solution a of the likelihood condition $\psi(t) = 0$ is a mean,—by Theorem I.

SECTION 5. SUMMARY

When the R. A. Fisher likelihood method is used to find an "optimum" scale for frequency functions, it sometimes happens that this scale is a well known mean or at least is a *substitutive* mean—See Equation (5). Or a simple transformation (15) may often put the frequency function into such a form. Conditions are given under which a scale will be a mean. Under further conditions this mean will be internal—at least as regards absolute values. Finally, under certain conditions, a scale will exist.

But for certain functions not satisfying these conditions, anomalies appear. The scale given by the usual likelihood condition may be a scale with a minimum likelihood. Sometimes the likelihood will be at neither a maximum nor a minimum. In certain simple cases, no scale exists. Furthermore, it may happen that the scales which are internal means have minimum likelihood and those that are external means have maximum likelihood. Among Pearson Types are found both anomalous functions and functions which would be regarded as regular as regards maximum likelihood.

In this problem of scaling, likelihood is proportional to *a posteriori* probability with the *a priori* probability taken as constant.

MOMENTS OF ANY RATIONAL INTEGRAL ISOBARIC SAMPLE MOMENT FUNCTION

By Paul S. Dwyer

Introduction

The problem of moments of moments has been investigated by a number of authors. The assumption of an infinite universe (or that of a finite universe with replacements) permits the application of the "algebraic" method, the method of semi-invariants as introduced by Thiele (1) and developed by C. C. Craig (2) and the combinatorial analysis method introduced by R. A. Fisher (3) and used by N. St. Georgescu (4). A combinatorial analysis method has the particular advantage that it enables one to compute separate terms of a given formula.

The formulae for moments of moments have been simplified through the use of new moment functions. Thiele introduced the half-invariant (1) which resulted in considerable condensation. More recently Prof. R. A. Fisher (3) has introduced the sample function k whose expected value is a half invariant. The most compact formulization presented thus far is his formulation of the half invariants of the sample k_r in terms of the half invariants of the universe. This very compactness, however, makes it difficult to compare results with those expressed in the more conventional sample functions. Dr. Wishart has written a paper (7) in which he shows, among other things, how the Fisher results can be translated to the more conventional (Craig) results and vice versa, but such translation is in general no simple matter. It appears that the Fisher results are not immediately useful to the statistician who desires the formulae to be expressed in terms of the usual sample moment function. On the other hand the Fisher formulization is a remarkable discovery toward that harmony which must be naturally inherent in the field of moments of moments. Soper (6, 111) expressed the general situation when he wrote, "If the terrifying overgrowth of algebraic formulation accompanying this branch of statistical inquiry is destined to have a chief utility in induction and going back to causes, then perhaps Dr. Fisher's way of estimating a sample will prove to be most fertile, but if it is to be applied to problems of deduction, say to problems of successive eventuation such as propagation, then Mr. Craig's plain moments seem to have a firmer hold on the exigencies of time."

It would appear then that the Fisher formulae and the Craig formulae are both needed. Georgescu (4) showed a partial connection between them in applying to the m functions a combinatory analysis somewhat similar to that applied by R. A. Fisher to the k function. It is the purpose of the present

paper to work out a combinatorial procedure for a more general sample function so that either the Fisher or Georgescu combinatorial results come out as special cases. In making such a generalization no limitation is placed on the sample function except that it be rational integral and that all terms are of the same weight. Thus the results are applicable to m_r, $m_r + k_r$, $m_r k_r$, etc. as well as to m_r and k_r although they are not applicable to $\sqrt{m_r}$ or $\frac{m_r}{k_r}$. In this way the important formulae for the moments of a new sample moment function will be available by simple substitution as soon as any such new function is defined by a rational integral isobaric expansion of power sums.

It is thus the purpose of this paper to determine the moments of a general moment function of the sample. This is done by keeping the multipliers of the various partitions of power sums indefinite until all manipulation is complete. It is then possible to assign the definite values of these multipliers which are associated with the desired sample function and to obtain the moment of the desired moment function in this way. Thus the Fisher result $\kappa(42)$, and the Craig result $S_{11}(\nu_4, \nu_2)$ are special cases of the new result $\lambda_{11}(f_4, f_2)$. It is obvious that it is not possible to carry the results using these general moment functions as far as Fisher and Wishart (3), (5), (7), have carried the results of the decidedly advantageous (from the standpoint of simplicity of result) k function and yet it is surprising to find the simplicity which can be obtained in the general case. Incidentally the introduction of the more general symbols clarifies the successive steps of the partition analysis which are somewhat confusing in any specific case because of the insertion of the value of the coefficients of the power sums in which the sample moment function is expressed.

This paper is divided into three parts. The first part includes the necessary definitions, the basic formulae, and the general development of the algebraic method. In order to facilitate the algebraic work there is inserted a table giving the expected values of all possible partition products of power sums whose weight $\leqq 8$. The second part deals with the different sample functions which might be used. The third part gives a list of the various partition formulae, of weight $\leqq 8$, which contain no unit parts and shows how these can be used in writing the chief variations of the formulae for moments of moments.

Part I

1. **General Moment Functions.** Different moment functions have been defined in various ways, but all moment functions have in common the property that they may be expressed in terms of the power sums. It appears sensible to use this expression in terms of power sums as the working algebraic definition of moment functions. For example the function k_3, which is defined by R. A. Fisher to be that function of the sample whose expected value is the third cumulant (half invariant) is to be given the working definition of

$$k_3 = \frac{n(3)}{(n-1)(n-2)} - \frac{3(2)(1)}{(n-1)(n-2)} + \frac{2(1)(1)(1)}{n(n-1)(n-2)}$$

where the numerical expressions in parentheses indicate power sums of the sample.

Every term in the definition of a sample function has a "weight" which is equal to the sum of the power sums whose product is indicated by the term. Thus the weight of each of the terms of k_3 is 3. If all the terms of a given moment function have the same weight, the function is called isobaric and the weight of the function is equal to the weight of each term. Thus k_3 is an isobaric moment function and its weight is 3. Since all the functions so far proposed are isobaric we limit this generalization of moment functions to isobaric moment functions although it is possible that a more complex analysis could be worked out for non-isobaric functions.

Generality demands the inclusion of every possible partition product of power sums. Such generality can be obtained by writing

$$
\begin{aligned}
f_1 &= a_1'(1) \\
f_2 &= a_2'(2) + a_{11}'(1)^2 \\
f_3 &= a_3'(3) + a_{21}'(2)(1) + a_{111}'(1)^3 \\
f_4 &= a_4'(4) + a_{31}'(3)(1) + a_{2^2}'(2)^2 + a_{21^2}'(2)(1)^2 + a_{1^4}(1)^4
\end{aligned}
$$

and in general

$$f_r = \sum a'_{p_1^{\pi_1} p_2^{\pi_2} p_s^{\pi_s}} (p_1)^{\pi_1} (p_2)^{\pi_2} \cdots (p_s)^{\pi_s}$$

where $(p_1)^{\pi_1} (p_2)^{\pi_2} \cdots (p_s)^{\pi_s}$ indicates any partition product of power sums, $a'_{p_1^{\pi_1} \cdots p_s^{\pi_s}}$ is its coefficient and the summation is taken for every possible partition. The number of parts of the partition is $\rho = \Sigma\pi$. It may be assumed, without loss of generality, that the partition is ordered, i.e.

$$p_1 \geqq p_2 \geqq p_3 \geqq \cdots \geqq p_s .$$

A natural numerical coefficient of each term is the number of ways the r units can be collected to form the given partition. This value is given by

$$\begin{pmatrix} 1^r \\ p_1^{\pi_1}\, p_2^{\pi_2} \cdots p_s^{\pi_s} \end{pmatrix} = \frac{r!}{(p_1!)^{\pi_1} (p_2!)^{\pi_2} \cdots (p_s!)^{\pi_s}\, \pi_1!\, \pi_2! \cdots \pi_s!}$$

If we set

$$a'_{p_1^{\pi_1} \cdots p_s^{\pi_s}} = \begin{pmatrix} 1^r \\ p_1^{\pi_1} \cdots p_s^{\pi_s} \end{pmatrix} a_{p_1^{\pi_1} \cdots p_s^{\pi_s}}$$

the definition of f_r becomes

$$f_r = \sum \begin{pmatrix} 1^r \\ p_1^{\pi_1} \cdots p_s^{\pi_s} \end{pmatrix} a_{p_1^{\pi_1} \cdots p_s^{\pi_s}} (p_1)^{\pi_1} \cdots (p_s)^{\pi_s}$$

In the present paper the capital letters are used to represent the corresponding

functions of the universe *as defined by the corresponding power sums of the universe.* Thus

$$F_r = \sum \binom{1^r}{p_1^{\pi_1} \cdots p_s^{\pi_s}} a_{p_1^{\pi_1} \cdots p_s^{\pi_s}} (P_1)^{\pi_1} \cdots (P_s)^{\pi_s}$$

represents the corresponding function of the universe. In the case of the moment about the mean and the semi-invariant the Greek letters μ and λ have been used to represent the corresponding function of the universe. In the case of functions whose notation is quite widely established, it is preferable to use the conventional notation, but in introducing new functions it appears wise to use the relationship between small and capital letters since the correspondence between the English and Greek alphabets is not exactly one to one. It should be particularly noticed that this notation does not agree with a previously accepted scheme of using the small English letter to indicate the function whose expected value is indicated by the corresponding Greek letter. In the present paper it is not the expected value property which serves as the basis of notation but rather the definition of the function in terms of the partition products of power sums.

2. The Working Definition of Moments About a Fixed Point. The sample functions defined by

$$m_1' = \frac{(1)}{n}, \qquad m_2' = \frac{(2)}{n}, \qquad m_3' = \frac{(3)}{n}, \cdots, m_r' = \frac{(r)}{n}$$

are obtained from f_r by placing

$$a_{p_1^{\pi_1} \cdots p_s^{\pi_s}} = \begin{cases} \dfrac{1}{n} \text{ when } s = 1, \pi_1 = 1, \text{ and } p_1 = r. \\ 0 \text{ in all other cases.} \end{cases}$$

The Greek μ' is used to indicate the corresponding function of the universe.

3. The Working Definition of Moments About the Mean. The moments about the mean are defined by

$$m_1' = \frac{(1)}{n}, \qquad m_2 = \frac{(2)}{n} - \frac{(1)(1)}{n^2},$$

$$m_3 = \frac{(3)}{n} - \frac{3(2)(1)}{n^2} + \frac{2(1)^3}{n^3}, \qquad m_4 = \frac{(4)}{n} - \frac{4(3)(1)}{n^2} + \frac{6(2)(1)^2}{n^3} - \frac{3(1)^4}{n^4}$$

and in general m_r is obtained from f_r by placing

$$a_{p_1^{\pi_1}\cdots p_s^{\pi_s}} = \begin{cases} \dfrac{1}{n} \text{ if } s = 1,\ \pi_1 = 1, \text{ and } p_1 = r. \\[2ex] \dfrac{(-1)^{\pi_2}}{(n)^{1+\pi_2}} \text{ if } p_1 > 1,\ \pi_1 = 1,\ s = 2, \text{ and } p_2 = 1. \\[2ex] \dfrac{(-1)^{r-1}(r-1)}{n^r} \text{ if } p_1 = 1,\ s = 1, \text{ and } \pi_1 = r. \\[2ex] 0 \text{ in all other cases.} \end{cases}$$

The corresponding moments of the universe are indicated by the conventional μ. For conciseness moments about the mean are referred to as "moments."

4. The Working Definition of the Half Invariants. The half invariant moment functions of Thiele, as applied to the sample power sums are [see C. C. Craig (2, 7–10) and Frisch (12, 20–21)]

$$l_1' = \frac{(1)}{n}, \qquad l_2 = \frac{(2)}{n} - \frac{(1)(1)}{n^2}, \qquad l_3 = \frac{(3)}{n} - \frac{3(2)(1)}{n^2} + \frac{2(1)^3}{n^3}$$

$$l_4 = \frac{(4)}{n} - \frac{4(3)(1)}{n^2} - \frac{3(2)^2}{n^2} + \frac{12(2)(1)^2}{n^3} - \frac{6(1)^4}{n^4}$$

and in general

$$l_r = \sum \frac{(-1)^{\rho-1}(\rho-1)!}{n^\rho} \begin{pmatrix} 1^r \\ p_1^{\pi_1} \cdots p_s^{\pi_s} \end{pmatrix} (p_1)^{\pi_1}(p_2)^{\pi_2} \cdots (p_s)^{\pi_s}$$

so that

$$a_{p_1^{\pi_1}\cdots p_s^{\pi_s}} = \frac{(-1)^{\rho-1}(\rho-1)!}{n^\rho}.$$

The corresponding moments of the universe are indicated, after Thiele (1) and Craig (2), by λ. R. A. Fisher (3) used κ while Georgescu (4) used s.

In the present paper these functions are referred to as "Thiele moments."

5. The k Functions of R. A. Fisher. The k statistics of R. A. Fisher are defined in terms of the sample power sums by

$$k_1' = \frac{(1)}{n}, \qquad k_2 = \frac{(2)}{n-1} - \frac{(1)^2}{n(n-1)},$$

$$k_3 = \frac{n(3)}{(n-1)(n-2)} - \frac{3(2)(1)}{(n-1)(n-2)} + \frac{2(1)^3}{n^{(3)}}$$

$$k_4 = \frac{n(n+1)(4)}{(n-1)^{(3)}} - \frac{4(n+1)(3)(1)}{(n-1)^{(3)}} - \frac{3(2)^2}{(n-2)^{(2)}} + \frac{12(2)(1)^2}{(n-1)^{(3)}} - \frac{6(1)^4}{n^{(4)}}.$$

These values and values for k_5 and k_6 are given by R. A. Fisher (3, 203–4) while algebraic methods of attaining them are presented in sections 16, 17. They are referred to as Fisher moments. The corresponding functions of the universe, if used, would be represented by K_r.

6. **The h Function.** Just as Fisher introduced a sample function whose expected value is a Thiele moment of the universe, so it is possible to introduce a function whose expected value is a moment of the universe. Such a function is defined by

$$h_1' = \frac{(1)}{n}, \qquad h_2 = \frac{(2)}{n-1} - \frac{(1)^2}{n(n-1)},$$

$$h_3 = \frac{n(3)}{(n-1)(n-2)} - \frac{3(2)(1)}{(n-1)(n-2)} + \frac{2(1)^3}{n^{(3)}}$$

$$h_4 \quad \frac{(n^2-2n+3)(4)}{(n-1)^{(3)}} - \frac{4(n^2-2n+3)(3)(1)}{n^{(4)}} - \frac{3(2n-3)(2)^2}{n^{(4)}} + \frac{6(2)(1)^2}{(n-1)^{(3)}} - \frac{3(1)^4}{n^{(4)}}.$$

Methods of obtaining the expansion of this function in terms of power sums are presented in section 18. The corresponding function of the universe, if it were used, would be represented by H_r.

7. **Other Moment Functions.** It is possible to obtain an indefinite number of moment functions. For example one might define a function of weight 2 whose variance equals μ_4, (or μ_2^2). It is possible by the methods of this paper to find expressions for such moments.

For reference purposes Table I is provided showing the values of a for each partition of weight <6 for the functions m', m, l, h, k. The values of

$$\binom{1^r}{p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s}}$$

are also inserted, in the left hand column, so that it is possible to read from the table the values for $f = m_r'$, m_r, l_r, k_r when $r < 6$.

8. **Products of f Functions.** The product of two or more isobaric functions is also isobaric and of weight equal to the sum of the weights of the functions. Thus

$$f_2 f_1 = [a_2(2) + a_{11}(1)(1)][a_1(1)] = a_2 a_1 (2)(1) + a_{11} a_1 (1)^2$$

$$f_2 f_1^2 = a_2 a_1^2 (2)(1)^2 + a_{11} a_1^2 (1)^4.$$

In multiplying f_{r_1} by f_{r_2} any term of f_{r_1} is of weight r_1 and when it is multiplied by any term of weight r_2, the result is a term of weight $r_1 + r_2$.

TABLE I

Coefficients of Products of Power Sums in the Expansion of Different Moment Functions

Numerical coefficient	a	m_r'	m_r	l_r	k_r	h_r
1	a_1	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$
1	a_2	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n-1}$	$\frac{1}{n-1}$
1	a_{11}	0	$\frac{-1}{n^2}$	$\frac{-1}{n^2}$	$\frac{-1}{n^{(2)}}$	$\frac{-1}{n^{(2)}}$
1	a_3	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{n}{(n-1)^{(2)}}$	$\frac{n}{(n-1)^{(2)}}$
3	a_{21}	0	$\frac{-1}{n^2}$	$\frac{-1}{n^2}$	$\frac{-1}{(n-1)^{(2)}}$	$\frac{-1}{(n-1)^{(2)}}$
1	a_{111}	0	$\frac{2}{n^3}$	$\frac{2}{n^3}$	$\frac{2}{n^{(3)}}$	$\frac{2}{n^{(3)}}$
1	a_4	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{n(n+1)}{(n-1)^{(3)}}$	$\frac{n^2-2n+3}{(n-1)^{(3)}}$
4	a_{31}	0	$\frac{-1}{n^2}$	$\frac{-1}{n^2}$	$-\frac{(n+1)}{(n-1)^{(3)}}$	$-\frac{n^2-2n+3}{n^{(4)}}$
3	a_{22}	0	0	$\frac{-1}{n^2}$	$\frac{-1}{(n-2)^{(2)}}$	$-\frac{2n-3}{n^{(4)}}$
6	a_{211}	0	$\frac{1}{n^3}$	$\frac{2}{n^3}$	$\frac{2}{(n-1)^{(3)}}$	$\frac{1}{(n-1)^{(3)}}$
1	a_{1111}	0	$\frac{-3}{n^4}$	$\frac{-6}{n^4}$	$\frac{-6}{n^{(4)}}$	$\frac{-3}{n^{(4)}}$
1	a_5	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{n^2(n+5)}{(n-1)^{(4)}}$	$\frac{n(n^2-5n+10)}{(n-1)^{(4)}}$
5	a_{41}	0	$\frac{-1}{n^2}$	$\frac{-1}{n^2}$	$-\frac{n(n+5)}{(n-1)^{(4)}}$	$-\frac{n^2-5n+10}{(n-1)^{(4)}}$
10	a_{32}	0	0	$\frac{-1}{n^2}$	$-\frac{n(n-1)}{(n-1)^{(4)}}$	$-\frac{n-2}{(n-1)^{(4)}}$

TABLE I—*Concluded*

Numerical coefficient	a	m'_r	m_r	l_r	k_r	h_r
10	a_{311}	0	$\frac{1}{n^3}$	$\frac{2}{n^3}$	$\frac{2(n+2)}{(n-1)^{(4)}}$	$\frac{n^2-4n+8}{n^{(5)}}$
15	a_{221}	0	0	$\frac{2}{n^3}$	$\frac{2(n-1)}{(n-1)^{(4)}}$	$+\frac{(2n-4)}{n^{(5)}}$
10	a_{2111}	0	$\frac{-1}{n^4}$	$\frac{-6}{n^4}$	$-\frac{6}{(n-1)^{(4)}}$	$-\frac{1}{(n-1)^{(4)}}$
1	a_{11111}	0	$\frac{4}{n^5}$	$\frac{24}{n^5}$	$\frac{24}{n^{(5)}}$	$\frac{4}{n^{(5)}}$

R. A. Fisher [3, 207] used the product $k_3^2k_2$ as an illustration of the algebraic method. The more general $f_3^2f_2$ gives

$$\begin{aligned} f_3^2f_2 &= [a_3(3) + 3a_{21}(2)(1) + a_{111}(1)^3]^2[a_2(2) + a_{11}(1)(1)] \\ &= a_3^2a_2(3)(3)(2) + a_3^2a_{11}(3)(3)(1)(1) + 6a_3a_{21}a_2(3)(2)^2(1) \\ &+ [6a_3a_{21}a_{11} + 2a_3a_2a_{111}](3)(2)(1)^3 + 9a_{21}^2a_2(2)^3(1)^2 + 2a_3a_{111}a_{11}(3)(1)^5 \\ &+ [6a_{21}a_{111}a_2 + 9a_{21}^2a_{11}](2)^2(1)^4 + [6a_{21}a_{111}a_{11} + a_2a_{111}^3](2)(1)^6 + a_{111}^3a_{11}(1)^8 \end{aligned}$$

which reduces to the value as given by him when the values of a are substituted from Table I.

9. The Expected Value of Any Partition Product. The expected values of partition products are well known and are indicated by

$$\begin{aligned} E(p_1) &= n\mu'_{p_1} \\ E(p_1)(p_2) &= n\mu'_{p_1+p_2} + n(n-1)\mu'_{p_1}\mu'_{p_2} \\ E(p_1)(p_2)(p_3) &= n\mu'_{p_1+p_2+p_3} + n(n-1)\,[\mu'_{p_1+p_2}\mu'_{p_3} + \mu'_{p_1+p_3}\mu'_{p_2} + \mu'_{p_2+p_3}\mu'_{p_1}] \\ &\qquad + n(n-1)\,(n-2)\,\mu'_{p_1}\mu'_{p_2}\mu'_{p_3}. \end{aligned}$$

and in general

$$E(p_1)^{\pi_1}(p_2)^{\pi_2}\cdots(p_s)^{\pi_s} = \sum n^{(\tau)}\begin{pmatrix} p_1^{\pi_1}\,p_2^{\pi_2}\cdots p_s^{\pi_s} \\ q_1^{\chi_1}\,q_2^{\chi_2}\cdots q_t^{\chi_t}\end{pmatrix}(\mu'_{q_1})^{\chi_1}(\mu'_{q_2})^{\chi_2}\cdots(\mu'_{q_t})^{\chi_t}$$

where $\tau = \chi_1 + \chi_2 + \chi_3 + \cdots + \chi_t$ and $\begin{pmatrix} p_1^{\pi_1}\,p_2^{\pi_2}\cdots p_s^{\pi_s} \\ q_1^{\chi_1}\,q_2^{\chi_2}\cdots q_t^{\chi_t}\end{pmatrix}$ indicates the

number of ways in which the partition $p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s}$ can be grouped to form the partition $q_1^{\chi_1} q_2^{\chi_2} \cdots q_t^{\chi_t}$.

The continued application of the result above leads to a large number of formulae. In order to make these results accessible I present in Table II the expected values of all partition products of weight ≤ 8. The essence of the table is the evaluation of the expression $\begin{pmatrix} p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s} \\ q_1^{\chi_1} q_2^{\chi_2} \cdots q_t^{\chi_t} \end{pmatrix}$. The numbers at the top of each column indicate the subscripts of the μ's which must, of course, be multiplied by $n^{(r)}$. The entries on the extreme left are the numerical coefficients associated with each row.

10. **The Expected Values of the f Functions.** With the use of Table II one is able to write expressions for the expected values of f_r when $r < 9$.

$$\mu_1'(f_1) = E(f_1) = a_1 n \mu_1'$$

$$\mu_1'(f_2) = E(f_2) = (a_2 + a_{11}) n \mu_2' + a_{11} n(n-1) \mu_1'^2$$

$$\mu_1'(f_3) = E(f_3) = (a_3 + 3a_{21} + a_{111}) n \mu_3' + 3(a_{21} + a_{111}) n(n-1) \mu_2' \mu_1'$$
$$+ a_{111} n(n-1)(n-2) \mu_1'^3 \text{ etc.}$$

If the expected values of the f functions are expressed in terms of the moments about the mean of the universe, these formulae become, since $\mu_1' = 0$

$$\mu_1'(f_1) = 0$$

$$\mu_1'(f_2) = (a_2 + a_{11}) n \mu_2$$

$$\mu_1'(f_3) = (a_3 + 3a_{21} + a_{111}) n \mu_3$$

$$\mu_1'(f_4) = (a_4 + 4a_{31} + 3a_{22} + 6a_{211} + a_{1111}) n \mu_4$$
$$+ 3(a_{22} + 2a_{211} + a_{1111}) n(n-1) \mu_2^2 \text{ etc.}$$

These may be written more symbolically as

$$\mu_1'(f_1) = 0$$

$$\mu_1'(f_2) = b_2 n \mu_2$$

$$\mu_1'(f_3) = b_3 n \mu_3$$

$$\mu_1'(f_4) = b_4 n \mu_4 + 3 b_{22} n(n-1) \mu_2^2 \text{ etc.}$$

11. **The Expected Value of Products of f Functions.** The expected value of products of f functions may be similarly found. For example

$$\mu_2'(f_2) = E(f_2^2) = E[a_2(2) + a_{11}(1)^2]^2 = a_2^2 E(2)^2 + 2a_2 a_{11} E(2)(1)(1) + a_{11}^2 E(1)^4.$$

TABLE II

Expected Values of Partition Products

weight = 1

coef.		n
	π	1
1	1	1

weight = 2

coef.		n	$n^{(2)}$
	π	2	11
1	2	1	
1	11	1	1

weight = 3

coef.		n	$n^{(2)}$	$n^{(3)}$
	π	3	21	111
1	3	1		
3	21	1	1	
1	1^3	1	3	1

weight = 4

coef.		n	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(4)}$
	π	4	31	22	211	1^4
1	4	1				
4	31	1	1			
3	22	1		1		
6	211	1	2	1	1	
1	1111	1	4	3	6	1

weight = 5

coef.		n	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(3)}$	$n^{(4)}$	$n^{(5)}$
	π	5	41	32	31^2	$2^2 1$	21^3	1^5
1	5	1						
5	41	1	1					
10	32	1		1				
10	311	1	2	1	1			
15	211	1	1	2		1		
10	2111	1	3	4	3	3	1	
1	1^5	1	5	10	10	15	10	1

TABLE II—*Continued*

weight = 6

coef.		n	$n^{(2)}$	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(4)}$	$n^{(4)}$	$n^{(5)}$	$n^{(6)}$
	π	6	51	42	33	411	321	2^3	31^3	$2^2\,1^2$	21^4	1^6
1	6	1										
6	51	1	1									
15	42	1		1								
10	33	1			1							
15	411	1	2	1		1						
60	321	1	1	1	1		1					
15	2^3	1		3				1				
20	31^3	1	3	3	1	3	3		1			
45	$2^2\,1^2$	1	2	3	2	1	4	1		1		
15	21^4	1	4	7	4	6	16	3	4	6	1	
1	1^6	1	6	15	10	15	60	15	20	45	15	1

weight

coef.		n	$n^{(2)}$	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(4)}$	$n^{(4)}$	$n^{(4)}$	$n^{(5)}$	$n^{(5)}$	$n^{(6)}$	$n^{(7)}$
	π	7	61	52	43	51^2	421	$3^2\,1$	32^2	41^3	321^2	$2^3\,1$	31^4	$2^2\,1^3$	21^5	1^7
1	7	1														
7	61	1	1													
21	52	1		1												
35	43	1			1											
21	51^2	1	2	1		1										
105	421	1	1	1	1		1									
70	$3^2\,1$	1	1		2			1								
105	32^2	1		2	1				1							
35	41^3	1	3	3	1	3	3			1						
210	321^2	1	2	2	3	1	2	2	1		1					
105	$2^3\,1$	1	1	3	3		3		3			1				
35	31^4	1	4	6	5	6	12	4	3	4	6		1			
105	$2^2\,1^3$	1	3	5	7	3	9	6	7	1	6	3		1		
21	21^5	1	5	11	15	10	35	20	25	10	40	15	5	10	1	
1	1^7	1	7	21	35	21	105	70	105	35	210	105	35	105	21	1

TABLE II—*Concluded*

weight = 8

$f_3^2 f_2$	coef.		n	$n^{(2)}$	$n^{(2)}$	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(4)}$	$n^{(4)}$	$n^{(4)}$	$n^{(4)}$	$n^{(4)}$	$n^{(5)}$	$n^{(5)}$	$n^{(5)}$	$n^{(6)}$	$n^{(6)}$	$n^{(7)}$	$n^{(8)}$
		π	8	71	62	53	44	611	521	431	42^2	$3^2\,2$	51^3	421^2	$3^2\,1^2$	$32^2\,1$	2^4	41^4	321^3	$2^3\,1^2$	31^5	$2^2\,1^4$	21^6	1^8
	1	8	1																					
	8	71	1	1																				
	28	62	1		1																			
	56	53	1			1																		
	35	44	1				1																	
	28	61^2	1	2	1			1																
	168	521	1	1	1	1			1															
	280	431	1	1		1	1			1														
	210	42^2	1		2		1				1													
$a_3^2 a_2$	280	$3^2\,2$	1		1	2						1												
	56	51^3	1	3	3	1		3	3				1											
	420	421^2	1	2	2	2	1	1	2	2	1			1										
$a_3^2 a_{11}$	280	$3^2\,1^2$	1	2	1	2	2	1		4		1			1									
$6a_3a_{21}a_2$	840	$32^2\,1$	1	1	2	3	1		2	1	1	2				1								
	105	2^4	1		4		3				6						1							
	70	41^4	1	4	6	4	1	6	12	4	3		4	6				1						
$6a_3a_{21}a_{11} + 2a_3a_{111}a_2$	560	321^3	1	3	4	5	3	3	6	9	3	4	1	3	3	3			1					
$9a_{21}^2 a_2$	420	$2^3\,1^2$	1	2	4	6	3	1	6	6	6	6		3		6	1			1				
$2a_3a_{111}a_{11}$	56	31^5	1	5	10	11	5	10	30	25	15	10	10	30	10	15		5	10		1			
$6a_{21}a_{111}a_2 + 9a_{21}^2 a_{11}$	210	$2^2\,1^4$	1	4	8	12	7	6	20	28	16	20	4	18	12	28	3	1	8	6		1		
$6a_{21}a_{111}a_{11} + a_{111}^2 a_2$	28	21^6	1	6	16	26	15	15	66	90	60	70	20	105	60	150	15	15	80	45	6	15	1	
$a_{111}^2 a_{11}$	1	1^8	1	8	28	56	35	28	168	280	210	280	56	420	280	840	105	70	560	420	56	210	28	1

Table II can now be used by indicating a_2^2 as a multiplier of $E(2)^2$, $2a_2a_{11}$ as a multiplier of $E(2)(1)(1)$ and a_{11}^2 as a multiplier of $(1)^4$. Then at once it is evident that

$$\begin{aligned}\mu_2'(f_2) &= (a_2^2 + 2a_2a_{11} + a_{11}^2)n\mu_4 + (a_2^2 + 2a_2a_{11} + 3a_{11}^2)n(n-1)\mu_2^2 \\ &= (a_2 + a_{11})^2 n\mu_4 + [(a_2 + a_{11})^2 + 2a_{11}^2]n(n-1)\mu_2^2 \\ &= b_2^2 n\mu_4 + (b_2^2 + 2b_{11}^2)n(n-1)\mu_2^2.\end{aligned}$$

Similarly

$$\mu_{11}'(f_3, f_2) = b_3b_2n\mu_5 + (b_3b_2 + 3b_{21}b_2 + 6b_{21}b_{11})n(n-1)\mu_3\mu_2$$

$$\begin{aligned}\mu_2'(f_3) = b_3^2n\mu_6 &+ (9b_{21}^2 + 6b_3b_{21})n(n-1)\mu_4\mu_2 + (b_3^2 + 9b_{21}^2)n(n-1)\mu_3^2 \\ &+ (9b_{21}^2 + 6b_{111}^2)n(n-1)(n-2)\mu_2^3\end{aligned}$$

etc.

where $b_3 = a_3 + 3a_{21} + a_{111}$, $b_{21} = a_{21} + a_{111}$, $b_{111} = a_{111}$. The important special cases are obtained by assigning the proper values to the a's as given in Table I. Thus

$$\mu_2'(m_2) = \frac{1}{n^3}[(n-1)^2\mu_4 + (n^2 - 2n + 3)(n-1)\mu_2^2]$$

which agrees with the corrected result of "Student" in 1908 (8, 3) and Tchouproff (10, 192). Similarly

$$\mu_{11}'(m_3, m_2) = \frac{1}{n^4}[(n-1)^2(n-2)\mu_5 + (n-1)(n-2)(n^2 - 5n + 10)\mu_3\mu_2]$$

$$\begin{aligned}\mu_2'(m_3) = \frac{1}{n^5}[&(n-1)^2(n-2)^2\mu_6 + (-6n + 15)(n-1)(n-2)^2\mu_4\mu_2 \\ &+ (n^2 - 2n + 10)(n-1)(n-2)^2\mu_3^2 + (9n^2 - 36n + 60)(n-1)(n-2)\mu_2^3]\end{aligned}$$

etc.

In the same way

$$\mu_2'(k_2) = \frac{\mu_4}{n} + \frac{(n^2 - 2n + 3)\mu_2^2}{n(n-1)}$$

$$\mu_{11}'(k_3, k_2) = \frac{\mu_5}{n} + \frac{(n^2 - 5n + 10)\mu_3\mu_2}{n(n-1)}$$

$$\mu_2'(k_3) = \frac{\mu_6}{n} + \frac{(-6n + 15)\mu_4\mu_2}{n(n-1)} + \frac{(n^2 - 2n + 10)\mu_3^2}{n(n-1)} + \frac{(9n^2 - 36n + 60)\mu_2^3}{n(n-1)(n-2)}$$

etc.

and

$$\mu_2'(m_2') = \frac{1}{n}[\mu_4 + (n-1)\mu_2^2]$$

$$\mu_{11}'(m_3', m_2') = \frac{1}{n}[\mu_5 + (n-1)\mu_3\mu_2]$$

$$\mu_2'(m_3') = \frac{1}{n}[\mu_6 + (n-1)\mu_3^2]$$

etc.

12. **The Expected Value of the Products of f Functions in Terms of the Thiele Moments of the Universe.** The formulae giving the μ's in term if the λ's are

$$\mu_2 = \lambda_2$$

$$\mu_3 = \lambda_3$$

$$\mu_4 = \lambda_4 + 3\lambda_2^2$$

$$\mu_5 = \lambda_5 + 10\lambda_3\lambda_2$$

$$\mu_6 = \lambda_6 + 15\lambda_4\lambda_2 + 10\lambda_3^2 + 15\lambda_2^3$$

$$\mu_r = \sum \binom{1^r}{p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s}} (\lambda_{p_1})^{\pi_1} (\lambda_{p_2})^{\pi_2} \cdots (\lambda_{p_s})^{\pi_s}$$

where the summation holds for those partitions having no unit parts. See the results of Craig (2, 7–11) and Frisch (12, 21). It is at once possible to express the moment formulae in terms of the Thiele moments of the universe. Thus the general results above become

$$\mu_2'(f_2) = b_2^2 n\lambda_4 + [3b_2^2 n + (b_2^2 + 2b_{11}^2)n(n-1)]\lambda_2^2$$

$$\mu_{11}'(f_3, f_2) = b_3b_2n\lambda_5 + [10b_3b_2n + (b_3b_2 + 3b_{21}b_2 + 6b_{21}b_{11})n(n-1)]\lambda_3\lambda_2$$

$$\begin{aligned}\mu_2'(f_3) = {} & b_3^2 n\lambda_6 + [15b_3^2 n + (9b_{21}^2 + 6b_3b_{21})n(n-1)]\lambda_4\lambda_2 \\ & + [10b_3^2 n + (b_3^2 + 9b_{21}^2)n(n-1)]\lambda_3^2 \\ & + [15b_3^2 n + (27b_{21}^2 + 18b_3b_{21})n(n-1) + (9b_{21}^2 + 6b_{111}^2)n(n-1)(n-2)]\lambda_2^3.\end{aligned}$$

13. **The Thiele Moments of the f's in terms of Thiele Moments.** It is now possible to reduce to the Thiele moments of the f's by means of the usual relations

$$\lambda_2(f_r) = \mu_2(f_r) - \mu_1'^2(f_r)$$

$$\lambda_{11}(f_{r_1}, f_{r_2}) = \mu_{11}'(f_{r_1}, f_{r_2}) - \mu_{10}'(f_{r_1}, f_{r_2})\mu_{01}'(f_{r_1}, f_{r_2})$$

$$\lambda_3(f_r) = \mu_3'(f_r) - 3\mu_2'(f_r)\mu_1'(f_r) + 2\mu_1'^3(f_r)$$

etc.

so that the results become

$$\lambda_2(f_2) = b_2^2 n\lambda_4 + 2[b_2^2 n + b_{11}^2 n(n-1)]\lambda_2^2$$

$$\lambda_{11}(f_3, f_2) = b_3 b_2 n\lambda_5 + \{3[b_3 b_2 n + b_{21} b_2 n(n-1)] + 6[b_3 b_2 n + b_{21} b_{11} n(n-1)]\}\lambda_3\lambda_2$$

$$\lambda_2(f_3) = b_3^2 n\lambda_6 + \{6[b_3^2 n + b_3 b_{21} n(n-1)] + 9[b_3^2 n + b_{21}^2 n(n-1)]\}\lambda_4\lambda_2$$
$$+ 9[b_3^2 n + b_{21}^2 n(n-1)]\lambda_3^2 + \{9[b_3^2 n + 2b_3 b_{21} n(n-1) + b_{21}^2 n(n-1) + b_{21}^2 n^{(3)}]$$
$$+ 6[b_3^2 n + 3b_{21}^2 n(n-1) + b_{111}^2 n(n-1)(n-2)]\}\lambda_2^3$$

etc.

The formulae as written are adapted to the partition representation of Part III.

When the f's are equal to the m's we have

$$\lambda_2(m_2) = \frac{(n-1)^2\lambda_4}{n^3} + \frac{2(n-1)\lambda_2^2}{n^2}.$$

$$\lambda_{11}(m_3, m_2) = \frac{(n-1)^2(n-2)\lambda_5}{n^4} + \frac{6(n-1)(n-2)\lambda_3\lambda_2}{n^3}$$

$$\lambda_2(m_3) = \frac{(n-1)^2(n-2)^2\lambda_6}{n^5} + \frac{9(n-1)(n-2)^2\lambda_4\lambda_2}{n^4}$$
$$+ \frac{9(n-1)(n-2)^2\lambda_3^2}{n^4} + \frac{6(n-1)(n-2)\lambda_2^3}{n^3}$$

etc.

which are the results as previously given by C. C. Craig (2, 55). In like manner when the $f_r = k_r$

$$\lambda_2(k_2) = \frac{\lambda_4}{n} + \frac{2\lambda_2^2}{n-1}$$

$$\lambda_{11}(k_3, k_2) = \frac{\lambda_5}{n} + \frac{6\,\lambda_3\lambda_2}{n-1}$$

$$\lambda_2(k_3) = \frac{\lambda_6}{n} + \frac{9\,\lambda_4\lambda_2}{n-1} + \frac{9\,\lambda_3^2}{n-1} + \frac{6n\,\lambda_2^3}{(n-1)(n-2)}$$

etc.

as given by R. A. Fisher [3, 210] while

$$\lambda_2(m_2') = \frac{1}{n}(\lambda_4 + 2\lambda_2^2)$$

$$\lambda_{11}(m_3', m_2') = \frac{1}{n}(\lambda_5 + 9\lambda_3\lambda_2)$$

$$\lambda_2(m_3') = \frac{1}{n}(\lambda_6 + 15\lambda_4\lambda_2 + 9\lambda_3^2 + 15\lambda_2^3).$$

etc.

14. Various Formulization of Results. Although different moment functions of the universe may be used it is customary to express the results in terms of universe moments about a fixed point, in terms of universe moments, or in terms of universe Thiele moments. It is possible to express results in any of the nine forms

$$\left.\begin{array}{r}\mu'(f_r)\\ \mu(f_r)\\ \lambda(f_r)\end{array}\right\}\text{ in terms of }\left\{\begin{array}{l}\text{moments about a fixed point }(\mu')\\ \text{moments }(\mu)\\ \text{Thiele moments }(\lambda)\end{array}\right.$$

where f_r represents the isobaric sample moment function of weight r. One purpose of such varied formulization is to discover the most compact form and also the one best adapted to use in the case of a normal universe or a universe whose moments obey some discoverable law. As suggested above Craig (2) has shown the relative compactness obtained by using $\lambda(m_r)$ and Thiele moments of the universe while R. A. Fisher (3) has shown the great additional compactness obtained by taking $f_r = k_r$.

15. The Application of the Algebraic Method to $\lambda_{21}(f_3, f_2)$. Before leaving the algebraic method it is perhaps wise to outline the steps in the case of a more involved problem. We take the example which R. A. Fisher (3, 207) has used in the case in which $f_r = k_r$. To find $\lambda_{21}(f_3, f_2)$.

The value of $f_3^2 f_2$ was found in section 8. To find its expected value it is only necessary to enter the coefficients of the different partition products in this expansion at the left of the corresponding rows as indicated in Table II.

The coefficient of any moment partition of the universe is found by multiplying each column entry by its corresponding left row entry and then by multiplying by $n^{(r)}$ as indicated at the top. Thus the coefficient of μ_8' is

$$(a_3^2a_2 + a_3^2a_{11} + 6a_3a_{21}a_2 + 6a_3a_{21}a_{11} + 2a_3a_{111}a_2 + 9a_{21}^2a_2 + 2a_3a_{111}a_{11} + 6a_2a_{21}a_{111}$$
$$+ 9a_{21}a_{21}a_{11} + 6a_{21}a_{111}a_{11} + a_{111}^2a_2 + a_{111}^2a_{11})n$$

which after some algebraic work reduces to

$$(a_3 + 3a_{21} + a_{111})^2(a_2 + a_{11})n = b_3^2b_2n.$$

In this manner it is possible to write the result either in terms of universe moments about a fixed point or in terms of universe moments. If moments are used, one may neglect all column partitions involving unity.

It should be noted that the a's defining k_r as given in Table I can be inserted here if desired. If these multipliers are introduced throughout the rows and columnar partitions involving unit parts are not used one will arrive at Table I of R. A. Fisher [3, 208] though there are some slight typographical errors in his rows for $(3)^2$ $(1)^2$ and (3) (2^2) (1).

Determining all the coefficients in this manner we find after considerable algebraic manipulation that

$$\mu'_{21}(f_3, f_2) = b_3^2 b_2 n\mu_8 + [b_3^2 b_2 + 9b_{21}^2 b_2 + 12b_3 b_{21} b_{11} + 6b_3 b_{21} b_2]n(n-1)\mu_6\mu_2$$
$$+ [2b_3^2 b_2 + 18b_{21}^2 b_2 + 18b_{21}^2 b_{11} + 6b_3 b_{21} b_2 + 12b_3 b_{21} b_{11}]n(n-1)\mu_5\mu_3$$
$$+ [2b_3^2 b_{11} + 9b_{21}^2 b_2 + 18b_{21}^2 b_{11} + 6b_3 b_{21} b_2]n(n-1)\mu_4^2 + [36b_{21}^2 b_2$$
$$+ 54b_{21}^2 b_{11} + 6b_3 b_{21} b_2 + 12b_3 b_{21} b_{11} + 12b_3 b_{111} b_{11} + 72b_{21} b_{111} b_{11}$$
$$+ 18b_{111}^2 b_2]n(n-1)(n-2)\mu_4\mu_2^2 + [b_3^2 b_2 + 6b_3 b_{21} b_2 + 12b_3 b_{21} b_{11}$$
$$+ 27b_{21}^2 b_2 + 90b_{21}^2 b_{11} + 36b_{21} b_{111} b_2 + 72b_{21} b_{111} b_{11} + 36b_{111}^2 b_{11}]n(n-1)(n-2)\mu_3^2\mu_2$$
$$+ [9b_{21}^2 b_2 + 18b_{21}^2 b_{11} + 36b_{21} b_{111} b_{11} + 6b_{111}^2 b_2 + 36b_{111}^2 b_{11}]n(n-1)(n-2)(n-3)\mu_2^4.$$

If $f_r = k_r$ the proper values of b are inserted and the expression above becomes that given by R. A. Fisher (3, 208). For example the coefficient of μ_2^4 is

$$\frac{(9n^3 - 63n^2 + 240n - 420)(n-3)}{n^2(n-1)^2(n-2)}$$

when

$$b_2 = \frac{1}{n}, \qquad b_3 = \frac{1}{n}, \qquad b_{11} = -\frac{1}{n(n-1)}$$

$$b_{21} = -\frac{1}{n(n-1)}, \qquad b_{111} = \frac{2}{n(n-1)(n-2)}.$$

The algebraic results involved in changing the general formula above to other functions are too extended to present here. A symbolic means of attaining them is included in later sections of the paper.

Part II. The Determination of Specific f Functions

16. **Functions Determined by the b's.** In Part I it was shown how various f functions are defined by giving definite values to the coefficients of the power sums. It is the purpose of this part of the paper to show how functions can be specified by means of their expected values in terms of moments of the universe. This is essentially the method used by R. A. Fisher in defining his k function and it is here extended to other functions. In this case the b's are first determined and the a's are then found from them. The first moments of f_1, f_2, f_3 were given in section 10. To these we add, as shown by Table II

$$\mu'_1(f_4) = (a_4 + 4a_{31} + 3a_{22} + 6a_{211} + a_{1111})n\mu'_4 + 4(a_{31} + 3a_{211} + a_{1111})n(n-1)\mu'_3\mu'_1$$
$$+ 3(a_{22} + 2a_{211} + a_{1111})n(n-1)\mu_2'^2 + 6(a_{211} + a_{1111})n(n-1)(n-2)\mu'_2\mu_1'^2$$
$$+ a_{1111}n(n-1)(n-2)(n-3)\mu_1'^4$$

etc.

These can be written more symbolically in terms of the b's

$$\mu_1'(f_1) = b_1 n\mu_1'$$

$$\mu_1'(f_2) = b_2 n\mu_2' + b_{11}n(n-1)\mu_1'^2$$

$$\mu_1'(f_3) = b_3 n\mu_3' + 3b_{21}n(n-1)\mu_2'\mu_1' + b_{111}n(n-1)(n-2)\mu_1'^3$$

$$\mu_1'(f_4) = b_4 n\mu_4' + 4b_{31}n(n-1)\mu_3'\mu_1' + 3b_{22}n(n-1)\mu_2'^2 + 6b_{211}n^{(3)}\mu_2'\mu_1'^2 + 6n^{(4)}\mu_1'^4,$$

and in general

$$\mu_1'(f_r) = \sum \begin{pmatrix} 1^r \\ p_1^{\pi_1}\ p_2^{\pi_2} \cdots p_s^{\pi_s} \end{pmatrix} b_{p_1^{\pi_1}\cdots p_s^{\pi_s}}\, n^{(\rho)}(\mu_{p_1}')^{\pi_1}(\mu_{p_2}')^{\pi_2} \cdots (\mu_{p_s}')^{\pi_s}.$$

The expansion of the function in terms of the power sums of the sample demands the determination of the a's. This can be accomplished by solving the equations

$$\begin{aligned}
a_1 &= b_1\\
a_2 + a_{11} &= b_2\\
a_{11} &= b_{11}\\
a_3 + 3a_{21} + a_{111} &= b_3\\
a_{21} + a_{111} &= b_{21}\\
a_{111} &= b_{111}\\
a_4 + 4a_{31} + 3a_{22} + 6a_{211} + a_{1111} &= b_4\\
a_{31} \qquad\quad + 3a_{211} + a_{1111} &= b_{31}\\
a_{22} + 2a_{211} + a_{1111} &= b_{22}
\end{aligned}$$

etc.

The solutions are

$$\begin{aligned}
a_1 &= b_1\\
a_2 &= b_2 - b_{11}\\
a_{11} &= b_{11}\\
a_3 &= b_3 - 3b_{21} + 2b_{111}\\
a_{21} &= b_{21} - b_{111}\\
a_{111} &= b_{111}\\
a_4 &= b_4 - 4b_{31} - 3b_{22} + 12b_{211} - 6b_{1111}\\
a_{31} &= b_{31} - 3b_{211} + 2b_{1111}\\
a_{22} &= b_{22} - 2b_{211} + b_{1111}\\
a_{211} &= b_{211} - b_{1111}\\
a_{1111} &= b_{1111}.
\end{aligned}$$

The values of a_r, at least for $r \leqq 4$, follow the law

$$a_r = \sum \binom{1^r}{p_1^{\pi_1} \cdots p_s^{\pi_s}} (-1)^{\rho-1} (\rho - 1)!\, b_{p_1^{\pi_1} \cdots p_s^{\pi_s}}$$

and

$a_{21} = \overline{|a_2 a_1|}$ where $\overline{|a_2 a_1|}$ indicates that $a_2 = b_2 - b_{11}$ is multiplied by $a_1 = b_1$, the rule of multiplication being suffixing of subscripts. Similarly $a_{22} = \overline{|a_2 a_2|} = \overline{|(b_2 - b_{11})(b_2 - b_{11})|} = b_{22} - 2b_{211} + b_{1111}$.

This statement illustrates a general theorem which will be established later in another paper by a different approach that for all cases

$$a_r = \sum \binom{1^r}{p_1^{\pi_1} \cdots p_s^{\pi_s}} (-1)^{\rho-1} (\rho - 1)!\, b_{p_1^{\pi_1} \cdots p_s^{\pi_s}}$$

and that

$$a_{r_1 \cdots r_t} = \overline{|a_{r_1} a_{r_2} \cdots a_{r_t}|}.$$

This theorem enables one to write, with comparative ease, the coefficient of any product of power sums in a sample function whose expected values is defined. For example the functional coefficient of (3)(2) in f_5 is

$$\overline{|a_3 a_2|} = \overline{|(b_3 - 3b_{21} + 2b_{111})(b_2 - b_{11})|} = b_{32} - b_{311} - 3b_{221} + 5b_{2111} - 2b_{11111}$$

while that of (3)(1)(1) is $\overline{|a_3 a_1 a_1|} = b_{311} - 3b_{2111} + 2b_{11111}$. If the expected value of the function is known the b's are determined and the values of the above expressions can be found by substitution.

17. **The Values of the Fisher Moments (k functions).** The k functions have been defined to be these functions whose expected values are the Thiele moments of the universe. Thus $\mu_1'(k_r) = \lambda_r$ and since

$$\lambda_r = \sum \binom{1^r}{p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s}} (-1)^{\rho-1} (\rho - 1)!\, (\mu'_{p_1})^{\pi_1} (\mu'_{p_2})^{\pi_2} \cdots (\mu'_{p_s})^{\pi_s}.$$

it follows at once that by comparison with $\mu_1'(f_r)$ in the last section, that

$$_{p_s^{\pi_s}} = \frac{(-1)^{\rho-1}(\rho - 1)!}{n^{(\rho)}}$$

Thus

$$b_1 = \frac{1}{n}; \quad b_2 = \frac{1}{n}; \quad b_{11} = -\frac{1}{n^{(2)}}; \quad b_3 = \frac{1}{n}; \quad b_{21} = \frac{-1}{n^{(2)}}; \quad b_{111} = \frac{2}{n^{(3)}};$$

$$b_4 = \frac{1}{n}; \quad b_{31} = \frac{-1}{n^{(2)}}; \quad b_{22} = \frac{-1}{n^{(2)}}; \quad b_{211} = \frac{2}{n^{(3)}}; \quad b_{1111} = \frac{-6}{n^{(4)}}; \text{ etc.}$$

The insertion of these values in the formulae of section 16 gives the values of a such as those indicated in Table I and in section 5. Thus the coefficient of (3)(2) in f_5 is

$$10(b_{32} - b_{311} - 3b_{221} + 5b_{2111} - 2b_{11111}) = -10\left[\frac{1}{n^{(2)}} + \frac{2}{n^{(3)}} + \frac{6}{n^{(3)}} + \frac{30}{n^{(4)}} + \frac{48}{n^{(5)}}\right]$$

$$- \frac{10n^{(2)}}{(n-1)^{(4)}}.$$

The coefficient of (3)(1)(1) is

$$10(b_{311} - 3b_{2111} + 2b_{11111}) = 10\left[\frac{2}{n^{(3)}} + \frac{18}{n^{(4)}} + \frac{48}{n^{(5)}}\right] = \frac{10(2n+4)}{(n-1)^{(4)}}.$$

18. **The h Functions.** It is also possible to define a function whose expected value is the moment of the universe. Thus $\mu_1'(h_r) = \mu_r$ where

$$\mu_r = \sum \binom{1^r}{p_1^{\pi_1} \cdots p_s^{\pi_s}} A_{p_1^{\pi_1} \cdots p_s^{\pi_s}} (\mu_{p_1}')^{\pi_1} (\mu_{p_2}')^{\pi_2} \cdots (\mu_{p_s}')^{\pi_s}$$

and

$$A_{p_1^{\pi_1} \cdots p_s^{\pi_s}} = \begin{cases} 1 \text{ if } s = 1,\ \pi_1 = 1, \text{ and } p_1 = r. \\ (-1)^{\pi_2} \text{ if } p_1 > 1,\ \pi_1 = 1,\ s = 2 \text{ and } p_2 = 1. \\ (-1)^{r-1}(r-1) \text{ if } p_1 = 1,\ s = 1, \text{ and } \pi_1 = r. \\ 0 \text{ in all other cases.} \end{cases}$$

Comparing with the value of $\mu_1'(f_r)$ in section 16 we have

$$b_{b_1^{\pi_1} \cdots b_s^{\pi_s}} = \frac{A_{p_1^{\pi_1} \cdots p_s^{\pi_s}}}{n^{(\rho)}}.$$

The substitution of these values of b in the results of section 16 gives the expansions of h_r in terms of power sums as illustrated by the formulae of section 6 and Table I. Thus the coefficient of (3)(2) is

$$10(b_{32} - b_{311} - 3b_{221} + 5b_{2111} - 2b_{11111})$$

$$= -10\left[0 + \frac{1}{n^{(3)}} + 0 + \frac{5}{n^{(4)}} + \frac{8}{n^{(5)}}\right] = \frac{-10(n-2)}{(n-1)^{(4)}}.$$

Similarly the coefficient of (3)(1)(1) in h_5 is

$$10(b_{311} - 3b_{2111} + 2b_{11111}) = 10\left[\frac{1}{n^{(3)}} + \frac{3}{n^{(4)}} + \frac{8}{n^{(5)}}\right] \quad \frac{10(n^2 - 4n + 8)}{n^{(5)}}$$

19. **The h' Functions.** One line of attack calls for the introduction of new moment functions which will result in simpler formulae. Thus for example,

C. C. Craig wrote (2, 37) "It rather seems that the best hopes of effectively further simplifying the problem of sampling for statistical characteristics lie either in the discovery of a new kind of symmetric function of all the observations which may be used to characterize frequency functions and which will be more amenable than either moments or semi-invariants for use in sampling problems, or in, what may very well prove to be much better and more feasible, the abandonment of the method of characterizing frequency functions by symmetric functions of all the observations altogether."

R. A. Fisher has shown that it is possible to introduce symmetric functions which do simplify the resulting formula appreciably. It is the purpose of this section to introduce an additional symmetric function which simplifies the resulting formulae to a much greater extent. It is admitted that this function does not have all the properties (such as invariance with respect to change of origin) possessed by the Thiele and Fisher functions, but it does not have the property of making the resulting formulae simple. It also has the advantage that $\mu(h_r') = \mu'(h_r')$.

The basic idea is to find a sample moment function whose expected value is 0. A first attempt, placing every $b = 0$, is of no avail since every a is also equal to 0 and there is no function. A second attempt is based on the idea of finding the function h whose expected value is $\mu_1'^r$. If the universe is assumed to be measured about its mean, as is conventional, it follows at once that $\mu_1' = 0$ and $\mu_1'(h_r) = 0$ so that

$$\mu_{\mu\nu}(h_{r_1}', h_{r_2}') = \mu_{\mu\nu}'(h_{r_1}', h_{r_2}').$$

This function then has the property that its moments about a fixed point and its moments are identical.

In order to discover its expansion in terms of power sums, we note

$$\mu_1'(h_r') = \mu_1'^r$$

and it follows at once by comparison with $\mu_1'(f_r)$ in section 16 that $b_{1^r} = \dfrac{1}{n^{(r)}}$ and $b_{p_1^{\pi_1} \cdots p_s^{\pi_s}} = 0$ in all other cases. The a's are determined in the usual way. Thus

$$a_2 = b_2 - b_{11} = -\frac{1}{n(n-1)}$$

$$a_{11} = b_{11} = \frac{1}{n(n-1)}$$

so that

$$h_2' = -\frac{1}{n(n-1)}[(2) - (1)(1)].$$

Similarly

$$h_3' = \frac{1}{n^{(3)}}[2(3) - 3(2)(1) + (1)^3]$$

$$h_4' = -\frac{1}{n^{(4)}}[6(4) - 8(3)(1) - 3(2)(2) + 6(2)(1)(1) - (1)^4]$$

and in general

$$h_r' = \frac{(-1)^{r-1}}{n^{(r)}}\left\{\sum (-1)^{\rho-1}[(p_1-1)!]^{\pi_1}[(p_2-1)!]^{\pi_2}\cdots[(p_s-1)!]^{\pi_s}\right.$$

$$\left.\binom{1^r}{p_1^{\pi_1}\cdots p_s^{\pi_s}}(p_1)^{\pi_1}\quad(p_s)^{\pi_s}\right\}.$$

In order to show the simple form in which results can be given we substitute the values of the b's in the results obtained above. Not only does $\mu_1'(h_r') = 0$, but by section 11

$$\lambda_2(h_2') = \mu_2(h_2') = \mu_2'(h_2') = \frac{2}{n(n-1)}\mu_2^2$$

$$\lambda_{11}(h_3', h_2') = \mu_{11}(h_3', h_2') = \mu_{11}'(h_3', h_2') = 0$$

$$\lambda_2(h_3') = \mu_2(h_3') = \mu_2'(h_3') = \frac{6}{n(n-1)(n-2)}\mu_2^3$$

while from section 15

$$\lambda_{21}(h_3', h_2') = \mu_{21}(h_3', h_2') = \mu_{21}'(h_3', h_2') = \frac{36\,\mu_3^2\mu_2}{n^2(n-1)^2(n-2)} + \frac{36(n-3)\,\mu_2^4}{n^2(n-1)^2(n-2)}.$$

It is to be noticed that these formulae contain very few terms and that the terms themselves involve very low moments of the universe. This simplicity has been attained without making any assumption such as normality, regarding the nature of the universe.

20. Table of Values of b for Different Functions When $r < 6$. This process of defining functions by means of expected values could be extended indefinitely. Perhaps it has been applied to enough functions to suggest the breadth of the applicability of the theory developed in Part I and Part III.

As the b's are the quantities which are used in the formulae I have provided Table III giving their values for the six functions, m_r', m_r, l_r, k_r, h_r, h_r' when $r = 1, 2, 3, 4, 5$. When the a's are known, the b's are computed from them according to the formulae of section 16.

TABLE III

Values of the b's for $r \leqq 5$

Num. coef.	b	m'_r	m_r	l_r	k_r	h_r	h'_r
1	b_1	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$
1	b_2	$\frac{1}{n}$	$\frac{n-1}{n^2}$	$\frac{n-1}{n^2}$	$\frac{1}{n}$	$\frac{1}{n}$	0
1	b_{11}	0	$-\frac{1}{n^2}$	$-\frac{1}{n^2}$	$-\frac{1}{n^{(2)}}$	$-\frac{1}{n^{(2)}}$	$\frac{1}{n^{(2)}}$
1	b_3	$\frac{1}{n}$	$\frac{(n-1)(n-2)}{n^3}$	$\frac{(n-1)(n-2)}{n^3}$	$\frac{1}{n}$	$\frac{1}{n}$	0
3	b_{21}	0	$-\frac{(n-2)}{n^3}$	$-\frac{n-2}{n^3}$	$-\frac{1}{n^{(2)}}$	$-\frac{1}{n^{(2)}}$	0
1	b_{111}	0	$\frac{2}{n^3}$	$\frac{2}{n^3}$	$\frac{2}{n^{(3)}}$	$\frac{2}{n^{(3)}}$	$\frac{1}{n^{(3)}}$
1	b_4	$\frac{1}{n}$	$\frac{(n-1)(n^2-3n+3)}{n^4}$	$\frac{(n-1)(n^2-6n+6)}{n^4}$	$\frac{1}{n}$	$\frac{1}{n}$	0
4	b_{31}	0	$-\frac{(n^2-3n+3)}{n^4}$	$-\frac{(n^2-6n+6)}{n^4}$	$-\frac{1}{n^{(2)}}$	$-\frac{1}{n^{(2)}}$	0
3	b_{22}	0	$\frac{2n-3}{n^4}$	$-\frac{(n^2-4n+6)}{n^4}$	$-\frac{1}{n^{(2)}}$	0	0
6	b_{211}	0	$\frac{n-3}{n^4}$	$\frac{2(n-3)}{n^4}$	$\frac{2}{n^{(3)}}$	$\frac{1}{n^{(3)}}$	0
1	b_{1111}	0	$-\frac{3}{n^4}$	$-\frac{6}{n^4}$	$-\frac{6}{n^{(4)}}$	$-\frac{3}{n^{(4)}}$	$\frac{1}{n^{(4)}}$
1	b_5	$\frac{1}{n}$	$\frac{(n-1)(n-2)(n^2-2n+2)}{n^5}$	$\frac{(n-1)(n-2)(n^2-12n+12)}{n^5}$	$\frac{1}{n}$	$\frac{1}{n}$	0
5	b_{41}	0	$-\frac{(n^3-4n^2+6n-4)}{n^5}$	$-\frac{(n^3-14n^2+36n-24)}{n^5}$	$-\frac{1}{n^{(2)}}$	$-\frac{1}{n^{(2)}}$	0
10	b_{32}	0	$\frac{n^2-4n+4}{n^5}$	$-\frac{(n^3-8n^2+24n-24)}{n^5}$	$-\frac{1}{n^{(2)}}$	0	0
10	b_{311}	0	$\frac{n^2-3n+4}{n^5}$	$\frac{2n^2-18n+24}{n^5}$	$\frac{2}{n^{(3)}}$	$\frac{1}{n^{(3)}}$	0
15	b_{221}	0	$-\frac{2(n-2)}{n^5}$	$\frac{2n^2-12n+24}{n^5}$	$\frac{2}{n^{(3)}}$	0	0
10	b_{2111}	0	$-\frac{n-4}{n^5}$	$-\frac{6(n-4)}{n^5}$	$-\frac{6}{n^{(4)}}$	$-\frac{1}{n^{(4)}}$	0
1	b_{11111}	0	$\frac{4}{n^5}$	$\frac{24}{n^5}$	$\frac{24}{n^{(5)}}$	$\frac{4}{n^{(5)}}$	$\frac{1}{n^{(5)}}$

Part III. Combinatory Methods

21. Partition Representation of Expected Value of f Functions. The formulae

$$\mu_1'(f_1) = b_1 n \mu_1'$$
$$\mu_1'(f_2) = b_2 n \mu_2' + b_{11} n(n-1)\mu_1'^2$$
$$\mu_1'(f_3) = b_3 n \mu_3' + 3b_{21} n(n-1)\mu_2'\mu_1' + b_{111} n(n-1)(n-2)\mu_1'^3$$
$$\mu_1'(f_4) = b_4 n \mu_4' + 4b_{31} n(n-1)\mu_2'\mu_1' + 3b_{22} n(n-1)\mu_2'^2$$
$$+ 6b_{211} n(n-1)(n-2)\mu_2'\mu_1'^2 + b_{1111} n^{(4)} \mu_1'^4$$

are "synthetically" given by the column partitions

$$\begin{array}{ccccc}
1 & & & & \\
 & & & & \\
2 & 1 & & & \\
 & 1 & & & \\
 & & & & \\
3 & 2 & 1 & & \\
 & 1 & 1 & & \\
 & & 1 & & \\
 & & & & \\
4 & 3 & 2 & 2 & 1 \\
 & 1 & 2 & 1 & 1 \\
 & & & 1 & 1 \\
 & & & & 1
\end{array}$$

The partition parts represent both the subscripts of the moments and the subscripts of the b's. If ρ indicates the number of parts, the n multiplier is $n^{(\rho)}$. The numerical coefficient is obtained by taking the sum of the entries in the column (the weight) and dividing it by the factorials of all entries times the factorials of all repeated entries as indicated by

$$\begin{pmatrix} 1^r \\ p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s} \end{pmatrix} = \frac{r!}{(p_1!)^{\pi_1}(p_2!)^{\pi_2}\cdots(p_s!)^{\pi_s}\,\pi_1!\,\pi_2!\cdots\pi_s!}.$$

The translation from the synthetic partition form to the expanded form is accelerated if the coefficients are known. These are provided in the following partition representation of the formula for $\mu_1'(f_r)$ when $r \leq 8$ and the results are expressed in terms of the moments of the universe

$$\begin{array}{ll}
\mu_1(f_1): & 0 \\
 & \\
\mu_1'(f_2): & 1 \\
 & 2 \\
 & \\
\mu_1'(f_3): & 1 \\
 & 3
\end{array}$$

$\mu_1'(f_4)$:

	1	**3**
	4	2
		2

$\mu_1'(f_5)$:

	1	**10**
	5	3
		2

$\mu_1'(f_6)$:

	1	**15**	**10**	**15**
	6	4	3	2
		2	3	2
				2

$\mu_1'(f_7)$:

	1	**21**	**35**	**105**
	7	5	4	3
		2	3	2
				2

$\mu_1'(f_8)$:

	1	**28**	**56**	**35**	**210**	**280**	**105**
	8	6	5	4	4	3	2
		2	3	4	2	3	2
					2	2	2
							2

The proper formula can be stated immediately from its synthetic representation. Thus for example

$$\mu_1'(f_6) = b_6 n\mu_6 + 15b_{42}n(n-1)\mu_4\mu_2 + 10b_{33}n(n-1)\mu_3^2 + 15b_{222}n(n-1)(n-2)\mu_2^3.$$

22. **Partition Representation of the Expected Value of a Product of f Functions.** Two column partitions may be used similarly to represent the expected values of the products of two f's, three column partitions for the expected value of the triple product, etc. In order to obtain all terms it is only necessary to combine every partition of each f in every possible way. The synthetic representation of $E(m_2, m_1)$ is

1	**1**	**2**	**1**
21	20	11	10
	01	10	10
			01

The sum of the entries in each row indicates the proper moment while the number of rows indicates the number of parts as in the preceding section. The n coefficient associated with a ρ rowed partition is then $n^{(\rho)}$. The b coefficient is indicated by the columnar entries. Thus

$$\mu_{11}'(f_2, f_1 = b_2 b_1 n\mu_3' + [b_2 b_1 + 2b_{11}b_1]n(n-1)\mu_2'\mu_1' + b_{11}b_1 n(n-1)(n-2)\mu_1'^3.$$

We verify this by the algebraic method

$$\begin{aligned}
\mu'_{11}(f_2, f_1) &= E\{[a_2(2) + a_{11}(1)(1)][a_1(1)]\} \\
&= E[a_2a_1(2)(1) + a_{11}a_1(1)^3] \\
&= a_2a_1[n\mu'_3 + n(n-1)\mu'_2\mu'_1 \\
&\qquad + a_{11}a_1[n\mu'_3 + 3n(n-1)\mu'_2\mu'_1 + n(n-1)(n-2)\mu'^3_1] \\
&= (a_2 + a_{11})a_1n\mu'_3 + (a_2 + a_{11})a_1n(n-1)\mu'_2\mu'_1 \\
&\qquad + 2a_{11}a_1\mu'_2\mu'_1 + a_{11}a_1n(n-1)(n-2)\mu'^3_1 \\
&= b_2b_1n\mu'_3 + b_2b_1n(n-1)\mu'_2\mu'_1 + 2b_{11}b_1n(n-1)\mu'_2\mu'_1 \\
&\qquad + b_{11}b_1n(n-1)(n-2)\mu'^3_1
\end{aligned}$$

as indicated.

It thus appears that the partition representation is a mnemonic device for indicating the solution as obtained by the algebraic method. A more formal justification is based upon the property that if

$$E(f_2) = b_2(2) + b_{11}(1)(1) \quad \text{and} \quad E(f_1) = b_1(1)$$

then $E(f_2, f_1)$ can be obtained by a symbolic multiplication of $b_2(2) + b_{11}(1)(1)$ by $b_1(1)$ where the b's are multiplied but the power sums are collected in all possible ways. Thus

$$E(f_2, f_1) = b_2b_1[(3) + (2)(1)] + b_{11}b_1[2(2)(1) + (1)^3]$$

which gives

$$E(f_2, f_1) = b_2b_1n\mu'_3 + b_2b_1n(n-1)\mu'_2\mu'_1 + 2b_{11}b_1n(n-1)\mu'_2\mu'_1 + b_{11}b_1n^{(3)}\mu'^3_1$$

as before.

This symbolic multiplication is generally true and serves as the real algebraic justification of the partition representation. It will be established in a later paper dealing with the more general case of a finite population. The general type of partition analysis has been used previously by Fisher (3) and Georgescu (4). Each has established it through analytic rather than algebraic means.

23. **Determination of the Coefficients.** Methods of determining the numerical coefficient have previously been given by such authors as Fisher (3), Wishart (5) (7) and Georgescu (4). If the f's are of different weight, the coefficients of any partition (an interchange of rows is not looked upon as changing the partition) is given by writing in the numerator the factorials of the different r's and in the denominator the factorials of all the different entries and the factorials of all repeated rows. Thus the coefficient of

$$\begin{matrix} 210 \\ 111 \\ 111 \end{matrix} \quad \text{is} \quad \frac{4!\,3!\,2!}{2!\,(1!)^7\,2!} = 72.$$

In case two or more functions have the same weight additional equivalent partitions are formed by interchange of columns. The reader is referred to the above papers for rules for determining the coefficients in the more involved cases though the coefficients are presented for all the two way partitions of the next section.

An alternative method of finding the coefficients is that given by C. C. Craig (2, 24–25) since it appears that the symbolic formulae used in the present paper are essentially his formulae for ν's in terms of λ's. For example his formula for ν_{44} (2, 22) is given symbolically by the formula for 44 in the next section. The only difference revealed is that the subscripts of the λ's are read by rows rather than by columns and that they are sometimes interchanged. The more precise formulization is needed for the present interpretation although it is not needed for Prof. Craig's purpose.

A third method utilizes the symbolic multiplication process stated in section 22. Subscripts of the b's are used to indicate which power sums are collected. Thus $[b_2(2) + b_{11}(1)(1)]^2$ gives

$$\underline{b_2b_2(4) + b_{20}b_{02}(2)(2)} + \underline{2[2b_{20}b_{11}(3)(1) + b_{200}b_{011}(2)(1)(1)]} + \underline{2b_{11}b_{11}(2)(2)}$$
$$\underline{+ 4b_{110}b_{101}(2)(1)(1) + b_{1100}b_{0011}(1)(1)(1)(1)}$$

where the underscored terms indicate the products given by $[b_2(2)]^2$, $2[b_2(2)]$ $[b_{11}(1)(1)]$, and $[b_{11}(1)(1)]^2$ respectively. This is represented by

1	1	4	2	2	4	1
22	20	21	20	11	11	10
	02	01	01	11	10	10
			01		01	01
						01
—	—			—		

The underscored terms are the only ones remaining when $\mu_1' = 0$.

This method is especially useful when a large number of formulae are to be computed, as in the next section.

24. The Partition Representation of Formulae of Total Weight $\leqq$ 8. The partition representation of $\mu_1'(f_r)$ when $r \leqq 8$ are given in section 21. The partition representation of the remaining formulae of total weight $\leqq 8$, which do not contain unit parts, are given below.

22	1	1	2
	22	20	11
		02	11

32	1	1	3	6
	32	30	12	21
		02	20	11

42	1	1	8	6	4	6	3	12
	42	40	31	22	30	21	20	20
		02	11	20	12	21	20	11
							02	11

33	1	6	9	1	9	9	6
	33	31	22	30	21	20	11
		02	11	03	12	11	11
						02	11

222	1	3	12	6	4	1	6	8
	222	220	211	201	111	200	200	110
		002	011	021	111	020	011	011
						002	011	101

52	1	1	10	10	5	10	20	10	20	15	60
	52	50	41	32	40	22	31	30	30	12	21
		02	11	20	12	30	21	20	11	20	20
								02	11	20	11

43	1	3	12	6	1	4	12	18	12	3	18	36	36
	43	41	32	23	40	13	31	22	30	03	21	12	21
		02	11	20	03	30	12	21	11	20	20	20	11
									02	20	02	11	11

322	1	2	4	12	3	1	4	6	12	12
	322	320	311	221	122	022	301	220	121	211
		002	011	101	200	300	021	102	201	111
	1	2	6	12	12	12	24	12	24	
	300	300	102	021	201	111	210	120	111	
	020	011	020	101	020	011	101	101	101	
	002	011	200	200	101	200	011	101	110	

62	1	1	12	15	6	30	20	15	20		
	62	60	51	42	50	41	32	40	31		
		02	11	20	12	21	30	22	31		
	15	30	120	45	10	60	120	90		15	90
	40	40	31	22	30	30	30	21		20	20
	20	11	20	20	30	12	21	21		20	20
	02	11	11	20	02	20	11	20		20	11
										02	11

53	**1**	**3**	**15**	**10**	**1**	**15**	**30**	**10**	**5**	**30**
	53	51	42	33	50	41	32	23	40	31
		02	11	20	03	12	21	30	13	22

	15	**60**	**90**	**15**	**30**	**10**	**30**	**60**	**90**	**90**	**45**	**60**
	40	31	22	13	31	30	30	30	12	21	20	20
	11	11	20	20	20	03	21	12	21	21	20	11
	02	11	11	20	02	20	02	11	20	11	11	11
											02	11

44	**1**	**12**	**16**	**8**	**48**	**1**	**16**	**18**
	44	42	33	41	32	40	31	22
		02	11	03	12	04	13	22

	6	**96**	**36**	**72**	**48**	**16**	**72**	**144**	**9**	**72**	**24**
	40	31	22	22	30	30	21	21	20	20	11
	02	11	20	11	12	03	21	12	20	11	11
	02	02	02	11	02	11	02	11	02	11	11
									02	02	11

422	**1**	**2**	**4**	**16**	**6**	**4**	**8**	**4**	**24**	**16**
	422	420	411	321	222	401	320	122	212	311
		002	011	101	200	021	102	300	210	111

	1	**16**	**6**	**12**
	400	310	220	211
	022	112	202	211

	1	**2**	**16**	**32**	**12**	**3**	**24**	**24**	**48**	**48**
	400	400	310	310	202	022	211	220	211	121
	020	011	110	101	200	200	200	101	101	200
	002	011	002	011	020	200	011	101	110	101

	8	**16**	**12**	**24**	**12**	**16**	**48**	**96**	**24**	**24**
	300	300	210	021	120	300	201	210	111	210
	120	021	210	201	102	111	120	111	111	201
	002	101	002	200	200	011	101	101	200	011

	3	**24**	**6**	**48**	**24**
	200	200	200	200	110
	200	110	200	110	110
	020	110	011	101	101
	002	002	011	011	101

332	1	1	9	12	6	2	18	18	6	12
	332	330	222	321	312	302	212	221	320	311
		002	110	011	020	030	120	111	012	021

	2	9	18	6
	301	220	211	310
	031	112	121	022

	9	18	6	12	12	18	9	72	18	36
	220	220	310	301	310	202	112	211	112	211
	110	101	020	020	011	110	200	110	110	101
	002	011	002	011	011	020	020	011	110	020

	1	6	12	9	18	36	36	18	36	72	36
	300	300	300	210	210	210	201	201	210	210	111
	030	012	021	120	102	012	111	021	101	111	111
	002	020	011	002	020	110	020	110	021	011	110

	9	18	36	6	36
	200	200	200	110	110
	110	101	110	110	110
	020	011	011	110	101
	002	020	011	002	011

2222	1	4	24	24	32	3	24	8
	2222	2220	2211	2201	2111	2200	2011	1111
		0002	0011	0021	0111	0022	0211	1111

	6	12	48	96	48
	2200	2200	2011	2011	1111
	0020	0011	0011	0101	1100
	0002	0011	0200	0110	0011

	24	48	96	16	48	16	32
	2001	2010	2100	0111	1011	1011	0111
	0201	0201	0111	0111	1110	0111	1101
	0020	0011	0011	2000	0101	1100	1010

	1	12	32	12	48
	2000	2000	2000	1100	1100
	0200	0200	0101	1100	0110
	0020	0011	0110	0011	0011
	0002	0011	0011	0011	1001

25. The Formulae for the Sample Moments about a Fixed Point in Terms of the Moments of the Universe. The partitions of section 21 and section 24 can be immediately interpreted to give the formulae for the moments of the sample function. For example .

$$\mu'_{11}(f_3, f_2) = b_3 b_2 n \mu_5 + (b_3 b_2 + 3b_{21} b_2 + 6b_{21} b_{11}) n(n-1)\mu_3\mu_2$$

and the value of $\mu'_{21}(f_3, f_2)$ as given in section 15 can be read by inspection. The value of the b's are to be inserted for any specific function. The coefficient of μ_2^3 in the expansion of $\mu'_3(f_2)$ is

$$(b_2^3 + 6b_2 b_{11}^2 + 8b_{11}^3) n(n-1)(n-2).$$

In case $f_2 = m_2$, $b_2 = \dfrac{n-1}{n^2}$, and $b_{11} = \dfrac{-1}{n^2}$ so that the coefficient is

$$\frac{(n-1)(n-2)(n^3 - 3n^2 + 9n - 15)}{n^5}$$

as indicated previously by Tchouproff (10, 192) and Church (9, 82).

The partitions of section 21 give the 8 formulae $\mu_{r,(N)}$ which Tchouproff gave (10, 155). In this case $f_r = m'_r$ and every b is 0 except those having single subscripts and these equal $\dfrac{1}{n}$.

The partitions of section 21 give the formulae $\nu_{r,(N)}$ which were given by Tchouproff (10, 186). In this case it is only necessary to take $f_r = m_r$ and to give the b's the proper values. Tchouproff has arranged his results according to decreasing powers of n. As an illustration we derive his result for $\nu_{4,(N)} = \mu'_1(m_4)$. From section 21

$$\mu'_1(f_4) = b_4 n\mu_4 + 3b_{22} n(n-1)\mu_2^2$$

and from Table II

$$b_4 = \frac{(n-1)(n^2 - 3n + 3)}{n^4} \quad \text{and} \quad b_{22} = \frac{2n-3}{n^4}$$

so that

$$\mu'_1(m_4) = \left(1 - \frac{4}{n} + \frac{6}{n^2} - \frac{3}{n^3}\right)\mu_4 + \left(\frac{6}{n} - \frac{15}{n^2} + \frac{9}{n^3}\right)\mu_2^2$$

$$= \mu_4 + \frac{1}{n}(6\mu_2^2 - 4\mu_4) - \frac{1}{n^2}(15\mu_2^2 - 6\mu_4) + \frac{1}{n^3}(9\mu_2^2 - 3\mu_4)$$

as indicated by him.

The partitions of section 24 also give formulae which have appeared before. For example the partitions

$$\begin{array}{ccc} 1 & 1 & 2 \\ 22 & 20 & 11 \\ & 02 & 11 \end{array}$$

which symbolize the formula

$$\mu_2'(f_2) = b_2^2 n\mu_4 + (b_2^2 + 2b_{11}^2)n(n-1)\mu_2^2$$

become

$$\mu_2'(m_2) = \frac{(n-1)}{n^3}\left[(n-1)\mu_4 + (n^2 - 2n + 3)\mu_2^2\right]$$

which was early derived by "Student" (8, 3) and Tchouproff (10, 192). Similarly the partitions of 222 and 2222 give the formula for and $\mu_3'(m_2)$ and $\mu_4'(m_2)$ which were given by Tchouproff (10, 192–193) and Church (9, 82).

Sections 21 and 24 can then be used to write the moments about a fixed point of a sample function in terms of the moments of the universe. In the case of new functions the b's must first be determined. Formulae involving unit columnar partitions are not included. If the formulae were desired in terms of moments about a fixed point of the universe, it would be necessary to write in addition all possible partitions. See for example the last formula of section 23.

26. The Formulae For Moments of Any Sample Function in Terms of Moments of the Universe. The partitions of sections 21 and 24 are also useful in writing the formulae for the moments of the sample moments. It is necessary to make the usual adjustments in changing from moments about a fixed point to moments:

$$\mu_2(f_r) = \mu_2'(f_r) - \mu_1'^2(f_r)$$

$$\mu_{11}(f_{r_1}, f_{r_2}) = \mu_{11}'(f_{r_1}, f_{r_2}) - \mu_{10}'(f_{r_1}, f_{r_2})\mu_{01}'(f_{r_1}, f_{r_2}).$$

The particular two way partitions which are involved in this adjustment are immediately recognizable. They are the ones which have an entry which is the only entry in the row and in the column in which it is. Thus $\begin{matrix}3\\220\\002\end{matrix}$ gives one of the terms contributing to $\mu_2'(f_2)\,\mu_1'(f_2)$. In addition its coefficient is the same, if sign is not considered, as the coefficient of $\mu_2'(f_2)\,\mu_1'(f_2)$ in the expansion of $\mu_3(f_2)$ in terms of moments of f_2. This has to be true since each is the number of ways of forming $\begin{matrix}220\\002\end{matrix}$. And so in general the remaining function of n accompanying this adjustment is the product of the coefficient associated with 22 and that associated with 2. The sign is plus when odd numbers of moments are multiplied and minus when even numbers of moments are multiplied. Hence $\begin{matrix}3\\220\\002\end{matrix}$ contributes $-3n^2b_2^3$ to the adjustment to moments and the total contribution of $\begin{matrix}3\\220\\002\end{matrix}$ to the value of $\mu_3(f_2)$ is $3b_2^3[n(n-1) - n^2] = -3b_2^3 n$. More

extensive study leads to the following general method of using the formulae of section 24.

A. Write the coefficient of every two way partition according to section 25.

B. Block off each single entry by drawing a line through its row and column. For example

$$6$$

The resulting partitions, 22, 2, 2 are called component parts.

C. Form new partitions by eliminating component parts one at a time, two at a time, three at a time, etc. from the original partition in all possible ways.

D. Form the coefficient of the resulting parts according to the methods of section 25. Multiply by $(-1)^{s-1}$ where s is the number of resulting parts. The values of b will not change.

E. Multiply in addition by $s - 1$ when the component parts are all taken separately.

As an example we find the contribution of the partition $\begin{array}{c}6\\2200\\0020\\0002\end{array}$ to the value of $\mu_4(f_2)$. It gives

$$6b_2^4[n(n-1)(n-2) - 3n^2(n-1) + 2n^3]\mu_4\mu_2\mu_2 = 12nb_2^4\mu_4\mu_2^2.$$

Similarly $\begin{array}{c}1\\2000\\0200\\0020\\0002\end{array}$ contributes

$$b_2^4[n^{(4)} - 4nn^{(3)} + 6n^3(n-1) - 3n^4]\mu_2^4 = 3b_2^4(n-2)\mu_2^4.$$

We use the method in finding the coefficient of μ_2^3 in the expansion of $\mu_3(m_2)$. We find first the coefficient of μ_2^3 in the expansion of $\mu_3(f_2)$. It is indicated by the partitions

$$\begin{array}{ccc}1 & 6 & 8\\ 200 & 200 & 110\\ 020 & 011 & 011\\ 002 & 011 & 101\end{array}$$

so that the coefficient of μ_2^3 is

$$\begin{aligned}b_2^3[n(n-1)(n-2) - 3n^2(n-1) + 2n^3] &+ 6b_2b_{11}^2[n(n-1)(n-2) - n^2(n-1)]\\ &+ 8b_{11}^3n(n-1)(n-2) = b_2^3(2n) + 6b_2b_{11}^2(-2n^2 + 2n)\\ &+ 8b_{11}^3n(n-1)(n-2).\end{aligned}$$

When $b_2 = \frac{n-1}{n^2}$ and $b_{11} = \frac{-1}{n^2}$ this becomes $\frac{2(n-1)(n^2-12n+15)}{n^5}$ as previously given by such authors as Tchouproff (10, 194), Church (9, 82), Carver (Richardson) (11, 271).

The general Tchouproff-Church formulae for the third and fourth moments of the variance may be written out in this way as may many other moment formulae which have not been printed.

27. The Thiele Moments of the Sample Function in Terms of the Moments of the Universe. It is possible also to write the Thiele moments of the sample function in terms of the moments of the universe. The technique is very similar to that of the previous section. The basis of the transformation is now the formula for Thiele moments in terms of moments about a fixed point rather than moments in terms of moments about a fixed point. The results are the same as those of the last section when a double or a triple product of f's is involved, but they differ with the introduction of a larger number of products. The partitions having component parts are broken up into these component parts as before but the parts are combined in all possible ways. Multipliers are determined as before with the exception that there is a multiplication by $(-1)^{s-1}(s-1)!$ where s is the number of resultant parts. Thus the term $\begin{matrix}2000\\0200\\0020\\0002\end{matrix}$ contributes $b_2^4[n^{(4)} - 4nn^{(3)} - 3n^2(n-1)^2 + 12n^3(n-1) - 6n^4]\mu_2^4 = -6b_2^4 n\mu_2^4$ to the value of $\lambda_4(f_2)$.

28. The Moments About a Fixed Point of the Sample Function in Terms of the Thiele Moments of the Universe. We return to the problem of section 25, only we wish to express the results in terms of the Thiele moments of the universe. We must use the formulae of section 12.

$$\mu_r = \sum \begin{pmatrix} 1^r \\ p_1^{\pi_1} \cdots p_s^{\pi_s} \end{pmatrix} (\lambda_{p_1})^{\pi_1} \cdots (\lambda_{p_s})^{\pi_s}$$

where $p_i \gtrless 1$.

Thus μ_r will contribute to all partitions of r and inversely the contributions to a given partition are composed only of these terms which are obtained by combining the different elements of the partition. Since the numerical coefficient in the expansion of μ_r is the number of ways in which the r units can be collected to form the partition, it follows at once that the complete λ coefficient can be obtained by grouping the parts of the partition in all possible ways, determining the coefficient of each according to the methods of section 25, and adding. In this way the formulae of section 21 can be used to give expansions in terms of partition moments. For example the representation of $\mu_1'(f_6)$

1	15	10	15
6	4	3	2
	2	3	2
			2

gives at once

$$b_6 n\lambda_6 + 15[b_6 n + b_{42} n(n-1)]\lambda_4\lambda_2 + 10[b_6 n + b_{33} n(n-1)]\lambda_3^2$$
$$+ 15[b_6 n + 3b_{42} n(n-1) + b_{222} n(n-1)(n-2)]\lambda_2^3.$$

The partitions of section 21 can be made to give the formula $\mu_1'(l_r)$ which were given by Thiele (1, 45–46). For example the formula for $\mu_1'(f_4)$ is indicated by

1	3
4	2
	2

so that

$$\mu_1'(f_4) = b_4 n\lambda_4 + 3[b_4 n + b_{22} n(n-1)]\lambda_2^2$$

and since

$$b_4 = \frac{(n-1)(n^2-6n+6)}{n^4} \quad \text{and} \quad b_{22} \quad \frac{2n-3}{n^4}$$

$$\mu_1'(l_4) = \frac{(n-1)(n^2-6n+6)\lambda_4}{n^3} - \frac{6(n-1)\lambda_2^2}{n^2},$$

which agrees with the result as given by him (1, 45).

The two way partitions of section 24 can be used similarly. This device for changing to the λ's is due to the ingenuity of R. A. Fisher who applied it to the case where $f_r = k_r$.

As an illustration we write from section 24 the value of $\mu_2'(f_2)$ in terms of λ's. The partition representation

1	1	2
22	20	11
	02	11

gives at once

$$b_2^2 n\lambda_4 + [b_2^2 n + b_2^2 n(n-1)]\lambda_2^2 + 2[b_2^2 n + b_{11}^2 n(n-1)]\lambda_2^2$$

which agrees with the result of section 12. The other illustrations of that section may be written out similarly.

As a final illustration of this technique we find the coefficient of λ_4^2 in the expansion of $\mu_{21}'(f_3, f_2)$. The partitions are

2	9	18	6
301	220	211	310
031	112	121	022

and the coefficient is

$$2[b_3^2b_2n + b_3^2b_{11}n(n-1)] + 9[b_3^2b_2n + b_{21}^2b_2n(n-1)]$$
$$+ 18[b_3^2b_2n + b_{21}^2b_{11}n(n-1)] + 6[b_3^2b_2n + b_3b_{21}b_2n(n-1)].$$

If the b's are inserted to form the k's, the first and last terms become 0 and the others give $\dfrac{27n - 45}{n(n-1)^2}$. This agrees with the value as given by R. A. Fisher (3, 208).

29. The Moments of the Sample Function in Terms of the Thiele Moments of the Universe. The partition representations of section 21 and section 24 can be used similarly to write formulae for the moments of the sample function in terms of the Thiele moments of the universe. It is only necessary to use the general plan of section 26, but to write the coefficient of every resulting partition according to the method of section 28. For example the partition

$$\begin{array}{c} 2000 \\ 0200 \\ 0020 \\ 0002 \end{array}$$

gives the coefficient

$$b_2^4[n + 4n^{(2)} + 3n^{(2)} + 6n^{(3)} + n^{(4)}] - 4b_2^4[n^2 + 3n^2(n-1) + n^2(n-1)(n-2)]$$
$$+ 6b_2^4[n^3 + n^3(n-1)] - 3b_2^4n^4 = b_2^4[n^4 - 4n^4 + 6n^4 - 3n^4] = 0.$$

30. The Thiele Moments of the Sample Function in Terms of the Thiele Moments of the Universe. The partition representations of section 21 and section 24 can also be interpreted to give the Thiele moments of the sample function in terms of the Thiele moments of the universe. The scheme is similar to that of section 29 except that the formulae for changing to Thiele moments are used as in section 27. For example the partition $\begin{array}{c} 2000 \\ 0200 \\ 0020 \\ 0002 \end{array}$ has now associated with it

$$b_2^4[n + 4n^{(2)} + 3n^{(2)} + 6n^{(3)} + n^{(4)}] - 4b_2^4[n^2 + 3n^2(n-1) + n^2(n-1)(n-2)]$$
$$- 3b_2^4n^2(n-1)^2 + 12b_2^4[n^3 + n^3(n-1)] - 6b_2^4n^4 = 0.$$

The application of this method enables one to write the formulae of section 13 (and others which they typify) with relative ease. It is now possible to complete the task left unfinished in section 15. We do not take the space necessary to write all the terms of $\lambda_{21}(f_3, f_2)$ since the lengthy expression can be obtained quite readily from the representation of section 24. One term, say the coefficient of $\lambda_6\lambda_2$, is represented by

1	9	12	6
330	222	321	312
002	110	011	020

and gives

$$9[b_3^2b_2n + b_{21}^2b_2n(n-1)] + 12[b_3^2b_2n + b_3b_{21}b_{11}n(n-1)] + 6[b_3^2b_2n + b_3b_{21}b_2n(n-1)]$$

which becomes $\dfrac{21}{n(n-1)}$ when $b_3 = b_2 = \dfrac{1}{n}$ and $b_{21} = b_{11} = \dfrac{-1}{n(n-1)}$. This agrees with the result given by R. A. Fisher (3, 209).

For simplicity of form it is logical to use this formulization of results, Thiele moments in terms of Thiele moments, and it has been used by Thiele (1), Craig (2), Fisher (3) and Georgescu (4). They however have used different sample moment functions. Thiele and Georgescu used the Thiele moments of the sample, Craig and Georgescu the moments while Fisher introduced the k function.

The present discussion deals with the corresponding partition moments of any rational integral isobaric moment function of the sample. The results indicated here give many of the results of the previous authors as special cases. For example the symbolic formula 44 of section 24 gives the $m^7\lambda_2(\mu_4)$ of Thiele (1, 45), the $S_{02}(\nu_2, \nu_4)$ of Craig (2, 57), the $\kappa(44)$ of R. A. Fisher (3, 210) as special cases when the formula 44 is given the interpretation of this section.

Some may prefer the Craig attack (2, 21–35) to the partition method. It should be noted that the formulae of sections 21 and 24 can be used in place of part of the Craig method. Thus his formulae (2, 22)

$$\nu_{80} = \lambda_{80} + 28\,\lambda_{60}\lambda_{20} + 56\,\lambda_{50}\lambda_{30} + \text{etc.}$$

$$\nu_{44} = \lambda_{44} + (12\,\lambda_{42}\lambda_{02} + 16\,\lambda_{33}\lambda_{11}) + \text{etc.}$$

are immediately obtainable from the symbolic formulae by writing λ's in place of b's and by using row, rather than column, subscripts. It is then necessary to compute the values of $\lambda_{k_1k_2}\ldots$ as given by him (2, 16–17, 40) and to insert in his expansions of $S_{kl}(\nu_m, \nu_n)$ in terms of ν's. For example

$$S_{11}(\nu_3, \nu_2) = \frac{1}{n}[\nu_{50} + (n-1)\nu_{32} - n\nu_{30}\nu_{02}] \qquad (2, 32)$$

and from the symbolic formulae of sections 21 and 24

$$\nu_{50} = \lambda_{50} + 10\lambda_{30}\lambda_{20}$$

$$\nu_{32} = \lambda_{32} + \lambda_{30}\lambda_{02} + 3\lambda_{12}\lambda_{20} + 6\lambda_{21}\lambda_{11}$$

$$\nu_{30} = \lambda_{30}$$

$$\nu_{20} = \lambda_{20}$$

so that

$$S_{11}(\nu_3\,, \nu_2) = \frac{1}{n}[\lambda_{50} + (n-1)\lambda_{32} + 9\lambda_{30}\lambda_{20} + (n-1)(6\lambda_{21}\lambda_{11} + 3\lambda_{21}\lambda_{20})] \qquad (2, 30)$$

which agrees with that given by Prof. Craig (aside from an obvious typographical error). The insertion of the values of λ gives the value as indicated by $\lambda_{11}(m_3\,, m_2)$ of section 13 and by the first method of the present section.

31. **Special Rules for the Determination of the Coefficients in the Case of the Fisher and Georgescu Analyses.** R. A. Fisher (3) gave a number of simple rules which assist greatly in the determination of the coefficients accompanying the partitions. Georgescu (4) also introduced special rules for the evaluation of the coefficients of the different partitions he used. It is not to be expected that all these rules are applicable in the more general case under present consideration, but the vanishing of such coefficients as that of $\begin{matrix}2000\\0200\\0020\\0002\end{matrix}$ leads one to suspect that there might be some rules which are applicable to this general case. A sensible method of procedure is to examine the rules of Fisher and Georgescu and determine if they hold in the more general analysis. The special rules of R. A. Fisher might be given somewhat as follows.

A. If a partition has a column with a single entry, that column may be eliminated and the factor n^{-1} introduced.

B. Any partition having a row with a single entry may be neglected.

C. "We may exclude any partition in which any set of rows is connected to its complementary set by a single column only."

D. In determining the algebraic coefficient of a partition the "pattern" is sufficient and precise entries are not needed. Thus the partitions $\begin{matrix}21\\11\end{matrix}$ and $\begin{matrix}35\\42\end{matrix}$, although they have different numerical factors, have associated with them the same function of n. This value is indicated by the pattern $\begin{matrix}xx\\xx\end{matrix}$ which has associated with it the function $\dfrac{1}{n-1}$. As a result of this property Fisher was able to provide a table (3, 223–226) of useful patterns which is of great assistance in writing the value of the coefficients.

E. Formulae of moments of k functions involving k_1 can be derived from corresponding formulae not involving k_1. "The effect upon the corresponding formula of adding a new unit part to the partition is (1) to modify every term in the formula by increasing the suffix of one of its κ functions by unity in every possible way, and (2) to divide the whole by n." (3, 206).

Two of the important Georgescu rules may be stated.

A′. The numerator function (aside from numerical coefficient) is not altered

if columns are changed to rows and vice versa. Thus the coefficient of s_3^3 in $S(3^2) = \dfrac{6N(N-1)}{(N+1)^3}$ and the coefficient of s_3^2 in $S(2^3)$ is $\dfrac{4N(N-1)}{(N+1)^4}$. Georgescu has replaced n by $N + 1$.

B′. All partitions which can be broken up into component parts have coefficients of 0. This is extended to include all partitions which have as component parts other partitions. Thus

$$\begin{matrix} 2100 \\ 1100 \\ 0012 \\ 0034 \end{matrix}$$

has a coefficient 0 as does the equivalent

$$\begin{matrix} 2010 \\ 1010 \\ 0102 \\ 0304 \end{matrix}$$

32. Special Rules for the Determination of the Coefficients in the More General Case. In the more general case we have

A. If a partition has a single column with a single entry, c, that column may be eliminated and the value b_c inserted as a multiplier. This is immediately evident since the contribution of that column to each term in the expansion is b_c times its value if the column were eliminated.

B. The coefficient of any partition having an entry which is the only entry in its row and column, is 0.

This rule, which saves considerable labor in that it makes unnecessary the computation of the coefficients of many of the partitions of section 24, is established in this way. Without loss of generality the partition may be represented by

$$\pi_{u+1,\,v+1} = \begin{matrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1v} & 0 \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2v} & 0 \\ c_{31} & c_{32} & c_{33} & \cdots & c_{3v} & 0 \\ \\ c_{u1} & c_{u2} & c_{u3} & \cdots & c_{uv} & 0 \\ 0 & 0 & 0 & & 0 & c_{u+1,\,v+1} \end{matrix}$$

and $\pi_{u,v}$ may represent the partition containing the first u rows and the first v columns. We determine the coefficient of $\pi_{u+1,\,v+1}$ in terms of the coefficient of $\pi_{u,v}$. Consider first any grouping of the u rows of $\pi_{u,v}$ into w rows. There will be w corresponding groupings of $\pi_{u+1,\,v+1}$ in which the last row is added, in turn, to each of the w rows and another $w + 1$ rowed term in which it is not

added. In each of the first w cases the coefficient by rule A is multiplied by $b_{c_{u+1}, v+1}$. In the case of the $w + 1$ rowed partition the coefficient is multiplied by $b_{c_{u+1}, v+1}$, and $n^{(w)}$ is replaced by $n^{(w+1)}$. A final adjustment takes care of the transition from the moment about a fixed point of the sample function to the Thiele moment of the sample function. This adjustment demands the multiplication of the coefficient of $\pi_{u,v}$ by $b_{c_{u+1}, v+1}\, n$ and the subtraction from the sum of the other terms. If B_w is the coefficient of the w rowed form, it follows at once that the corresponding coefficient is

$$B_w b_{c+1, v+1} [wn^{(w)} + n^{(w+1)} - n\, n^{(w)}] = 0.$$

This holds for the expansion of any term of $\pi_{u,v}$ and hence the coefficient of $\pi_{u+1, v+1}$ is 0. Of course the argument holds if the partition has more than 2 component parts.

It thus appears that this rule holds not only for k_r and m_r as Fisher and Georgescu have noted, but for f_r.

C. The coefficient of any partition which can be broken into component parts is 0. In this sense a component part is any group of rows or columns which have no entry in common with any other group of rows or columns. It corresponds in matrix language to a matrix which results when one matrix is zero bordered by another matrix although rows and columns may thereafter be interchanged.

The proof of this more general case follows the general line of the simpler case although the reasoning is more complicated. For example the coefficient of

$$\begin{matrix} c_{11} & c_{12} \cdots c_{1v} & 0 & 0 \\ c_{21} & c_{22} \cdots c_{2v} & 0 & 0 \\ c_{31} & c_{32} \cdots c_{3v} & 0 & 0 \\ \\ c_{u1} & c_{u2} \cdots c_{uv} & 0 & 0 \\ 0 & 0 \quad \cdots 0 & c_{u+1, v+1} & c_{u+1, v+2} \\ 0 & 0 \quad \cdots 0 & c_{u+2, v+1} & c_{u+2, v+2} \end{matrix}$$

is 0 since any w rowed term of the $\pi_{u,v}$ contributes

$$\begin{aligned} B_w b_{c_{u+1}, v+1 + c_{u+2}, v+1}\, b_{c_{u+2}, v+2 + c_{u+2}, v+2} &[wn^{(w)} + n^{(w+1)} - .nn^{(w)}] \\ + B_w b_{c_{u+1}, v+1 c_{u+2}, v+2}\, b_{c_{u+1}, v+2 c_{u+2}, v+2} &[w(w-1)\, n^{(w)} + 2wn^{(w+1)} + n^{(w+2)} \\ &- n(n-1)\, n^{(w)}] = 0. \end{aligned}$$

Other special rules of Fisher and Georgescu do not hold in the general case. Thus Fisher rule B is not generally true since the partitions

$$\begin{matrix} 12 \\ 30 \end{matrix} \quad \text{and} \quad \begin{matrix} 22 \\ 20 \end{matrix}$$

have respective algebraic coefficients of $b_4b_2n + b_{31}b_2n(n - 1)$ and

$$b_4b_2n + b_{22}b_2n(n - 1)$$

and these are not in general equal to 0.

The Fisher rule C is replaced by the somewhat less general C of the present section.

The Fisher rule D is not applicable in the general case. The Fisher rule D is applicable in all cases in which the value of the $b_{p_1^{\pi_1} \cdots p_s^{\pi_s}}$ is completely determined by the number of parts for in this case the particular value of each part is not pertinent. We may say then that the Fisher rule D is applicable to all cases in which $b_{p_1^{\pi_1} p_2^{\pi_2} \cdots p_s^{\pi_s}}$ is a function of ρ, n where ρ is the number of parts. This condition is satisfied by $b_{p_1^{\pi_1} \cdots p_s^{\pi_s}} = \dfrac{(-1)^{\rho}\,(\rho - 1)!}{n^{(\rho)}}$ and the coefficients are worked out for it in Fisher's paper. The same method is applicable to other functions satisfying the general condition although the values of the coefficients will of course vary with the definition of b.

The Fisher rule E is not applicable to the general case. Its validity, from an algebraic standpoint, depends upon the Fisher property B which is not generally applicable. The Fisher rule E as applied to the more general case gives correct terms but it does not give all the terms. For example the Fisher rule E applied to $\lambda_2(k_2)$ gives

$$\lambda_2(k_2) = \frac{\lambda_4}{n} + \frac{2\lambda_2^2}{n - 1}$$

$$\lambda_{21}(k_2, k_1) = \frac{\lambda_5}{n^2} + \frac{4\lambda_3\lambda_2}{n(n - 1)}.$$

The application of a corresponding rule to

$$\lambda_2(f_2) = b_2^2 n\lambda_4 + 2[b_2^2 n + b_{11}n(n - 1)]\lambda_2^2$$

would give

$$\lambda_{21}(f_2, f_1) = b_2^2 b_1 n\lambda_5 + 4[b_2^2 b_1 n + b_{11}^2 b_1 n(n - 1)]\lambda_3\lambda_2$$

while the correct result is indicated by

$$\begin{array}{cccc}
1 & 4 & 2 & 4 \\
221 & 210 & 201 & 111 \\
 & 011 & 020 & 110
\end{array}$$

and is

$$\lambda_{21}(f_2 f_1) = b_2^2 b_1 n\lambda_5 + 4[b_2^2 b_1 n + b_2 b_{11} b_1 n(n-1)]\lambda_3\lambda_2 + 2[b_2^2 b_1 n + b_2^2 b_1 n(n - 1)]\lambda_3\lambda_2$$
$$+ 4[b_2^2 b_1 n + b_{11}^2 b_1 n(n - 1)]\lambda_3\lambda_2.$$

The difference is due to the vanishing of the two middle terms in the case of the k functions.

The rule B′, which Georgescu found most useful in computing and checking his formulae, is not generally true. It is not even true in the case of the k function, as can be discovered by using it on the list given by R. A. Fisher (3, 210). It is interesting to note that the Georgescu method, while not being able to utilize many of the special rules of the Fisher method, does use this rule which is not in general adaptable to the Fisher method.

33. **Special Rules in the Case of the h' Functions.** Special rules can be worked out for other sample functions. As an illustration we examine the function h'_r which was defined in section 19. It is recalled that $b_{1^\rho} = \frac{1}{n^{(\rho)}}$ and that $b_{p_1^{\pi_1} \cdots p_s^{\pi_s}} = 0$ for all other cases. It follows at once that

A. Any partition having any entry other than unity (or zero) may be neglected.

B. The value of b_{1^ρ} is $\frac{1}{n^{(\rho)}}$.

As an illustration we write the value $\lambda_{21}(h'_3, h'_2)$. From the partitions of section 24 we select

36		**36**
111		110
111	and	110
110		101
		011

as being the only partitions making a contribution. The result of section 19 follows at once.

34. **The Case of a Normal Universe.** A normal universe is characterized by the relationship that $\lambda_r = 0$ when $r > 2$. It follows that it is only necessary to compute the coefficients of those partitions giving powers of λ_2.

Wishart (5) (7) has developed the partition analysis of the k function in the case of a normal parent while Georgescu has studied the corresponding m function. It is not the purpose of this section to make extensive study of the case of the normal parent but simply to indicate that the results of section 24 are immediately applicable. As an illustration we write the values of $\lambda_1(f_2)$, $\lambda_2(f_2)$, $\lambda_3(f_2)$ and $\lambda_4(f_2)$ in the case of a normal universe. The terms are given successively, by

1	**2**	**8**	**48**
2	11	110	1100
	11	011	0110
		101	0011
			1001

and hence

$$\lambda_1(f_2) = b_2 n\lambda_2$$

$$\lambda_2(f_2) = 2[b_2^2 n + b_{11}^2 n(n-1)]\lambda_2^2$$

$$\lambda_3(f_2) = 8[b_2^3 n + 3b_2 b_{11}^2 n(n-1) + b_{11}^3 n(n-1)(n-2)]\lambda_2^3$$

$$\lambda_4(f_2) = 48[b_2^4 n + 6b_2^2 b_{11}^2 n(n-1) + b_{11}^4 n(n-1) + 4b_2 b_{11}^3 n(n-1)(n-2) + 2b_{11}^4 n(n-1)(n-2) + b_{11}^4 n(n-1)(n-2)(n-3)]\lambda_2^4.$$

It is only necessary to substitute the b's to obtain the results for different values of f. This is done in Table IV.

TABLE IV

The first four Thiele moments of f_2 for various sample functions in the case of a normal universe

Sample function	$\lambda_1(f_2)$	$\lambda_2(f_2)$	$\lambda_3(f_2)$	$\lambda_4(f_2)$
m_2	$\frac{(n-1)}{n}\lambda_2$	$\frac{2(n-1)}{n^2}\lambda_2^2$	$\frac{8(n-1)}{n^3}\lambda_2^3$	$\frac{48(n-1)\lambda_2^4}{n^4}$
k_2	λ_2	$\frac{2\lambda_2^2}{n-1}$	$\frac{8\lambda_2^3}{(n-1)^2}$	$\frac{48\lambda_2^4}{(n-1)^3}$
l_2	$\frac{(n-1)}{n}\lambda_2$	$\frac{2(n-1)}{n^2}\lambda_2^2$	$\frac{8(n-1)\lambda_2^3}{n^3}$	$\frac{48(n-1)\lambda_2^4}{n^4}$
m_2'	λ_2	$\frac{2\lambda_2^2}{n}$	$\frac{8\lambda_2^3}{n^3}$	$\frac{48\lambda_2^4}{n^4}$
h_2	λ_2	$\frac{2\lambda_2^2}{n-1}$	$\frac{8\lambda_2^3}{(n-1)^2}$	$\frac{48\lambda_2^4}{(n-1)^3}$
h_2'	0	$\frac{2\lambda_2^2}{n(n-1)}$	$\frac{8(n-2)\lambda_2^3}{n^2(n-1)^2}$	$\frac{48(n^2-3n+3)\lambda_2^4}{n^3(n-1)^3}$

One surmises that the general value of

$$\lambda_r(f_2) \text{ is } 2^{r-1}(r-1)!\,\lambda_2^r B: \begin{array}{l} 11000\cdots 0 \\ 01100\cdots 0 \\ 00110\cdots 0 \\ \\ 00000\cdots 11 \\ 10000\cdots 01 \end{array}$$

where B represents the b coefficient of the r rowed partition. This induction appears consistent with the fact that

$$\lambda_{r+1}(k_2) = \frac{2^r r! \lambda_2^{r+1}}{(n-1)^r}$$

as shown by John Wishart (7). The whole subject of the Thiele moments of the general function in the case of a normal universe would make an interesting subject of investigation.

35. Summary and Conclusion. The contributions of this paper include

1. The definitions of specific moment functions in terms of power sums.
2. The use of indeterminate multipliers in representing a general isobaric moment function.
3. The finding of the expected value of products of these functions by algebraic methods.
4. The use of tables in writing these expected values in terms of moments (or of moments about a fixed point) of the universe.
5. The finding of the expected values of specific moment functions by substitution.
6. Means of establishing the expansion of new moment functions which are defined by their expected values.
7. The introduction of the sample function of weight r whose expected value is μ_r.
8. The introduction of the sample function of weight r whose expected value is $\mu_1'^r$.
9. The two way partition formulae of weight ≤ 8 which do not involve unit parts.

The use of these partition formulae in writing:

10. The moments about a fixed point of f_r in terms of moments.
11. The moments of f_r in terms of moments.
12. The Thiele moments of f_r in terms of moments.
13. The moments about a fixed point of f_r in terms of Thiele moments.
14. The moments of f_r in terms of Thiele moments.
15. The Thiele moments of f_r in terms of Thiele moments.
16. Special rules in the case of Thiele moments.
17. The applicability of these results to a given sample moment function and hence the derivation of varied results, of such authors as Thiele, Tchouproff, Church, Fisher, Craig, and Georgescu, from the same partition formulae.
18. The simplicity of the formulae when h_r' is used as the sample function.
19. The application of the synthetic formulae to the Craig method.
20. The applicability of the theory to a normal universe.

The introduction of such general procedure opens up a wide field for future study. It is impossible in a single paper dealing with so broad a subject to do more than to outline the general scheme by which two way partitions can be

used as a central formulization of the various formulae for moments of moments. More detailed proofs and more extensive analysis of the more important of the special cases will undoubtedly be supplied by later writers.

In later papers the author will show how the partition representation can be used in the case of multivariate distributions and how it can also be used, in connection with the sampling polynomials introduced by H. C. Carver (11), to represent the more complex formulae obtained in the case of finite sampling.

It is obvious that the author is indebted to the classical moment studies of Fisher and Craig. He also wishes to acknowledge his indebtedness to Prof. Craig and to Prof. Carver who have read the manuscript and have made valuable suggestions.

THE UNIVERSITY OF MICHIGAN.

REFERENCES

(1) THIELE, T. N. "Theory of observations." London 1903. C. and E. Layton. Reprinted in "Annals of Mathematical Statistics," 2 (1931), pp. 165–306.

(2) CRAIG, C. C. "An application of Thiele's semi-invariants to the sampling problem." Metron, 7 (1928–29), pp. 3–74.

(3) FISHER, R. A. "Moments and product moments of sampling distributions." Proceedings of the London Mathematical Society. Series 2, vol. 30 (1930), pp. 199–238.

(4) ST. GEORGESCU, N. "Further contributions to the sampling problem." Biometrika, 24 (1932), pp. 65–107.

(5) WISHART, JOHN. "The derivation of certain high order sampling product moments from a normal population." Biometrika, 22 (1930), pp. 224–238.

(6) SOPER, H. E. "Sampling moments of moments of samples of n units each drawn from an unchanging sampled population, from the point of view of semi-invariants." Journal of Royal Statistical Society, 93 (1930), pp. 104–114.

(7) WISHART, JOHN. "A comparison of the semi-invariants of the distributions of moment and semi-invariant estimates in samples from an infinite population." Biometrika, 25 (1932), pp. 52–60.

(8) "Student." "The probable error of the mean." Biometrika, 6 (1908), pp. 1–25.

(9) CHURCH, A. E. R. "On the moments of the distribution of squared standard deviations for samples of n drawn from an indefinitely large population." Biometrika, 17 (1925), pp. 79–83.

(10) TCHOUPROFF, A. A. "On the mathematical expectation of the moments of frequency distributions." Biometrika, 12 (1918–19), pp. 140–169, 185–210.

(11) CARVER, H. C. "Fundamentals of sampling." Editorial. "Annals of Mathematical Statistics," 1 (1930), pp. 101–121; 260–274.

(12) FRISCH, RAGNAR. "Sur les semi-invariants et moments employes dans l'etude des distributions statistiques." Skrifter utgitt av det Norske Videnskaps-Akademi i Oslo. II Hist.-Filos Klasse (1926), no. 3.

NOTES

A COEFFICIENT OF CORRELATION BETWEEN SCHOLARSHIP AND SALARIES

INTRODUCTION

Some might doubt that it is correct to apply a coefficient of correlation to show the relationship between scholarship and salaries. This coefficient can be trusted to give at least a rough approximation, which is all that is necessary in the inexact science of vocation. It is fictitious accuracy to be too finical in the application of formulas. Therefore, a coefficient of correlation between scholarship and salaries is a valuable part of human knowledge.

Would it be worth while to find this coefficient if it is based upon the experience of the American Telegraph and Telephone Company? Since the employment practices of this company are not representative of the employment practices of business at large, one might doubt the validity of drawing general conclusions from such specialized data. The coefficient for business at large is probably less than the coefficient for the Bell System; the value of this knowledge is enhanced if we know the latter coefficient. Since this company is very large, a coefficient between scholarship and salaries would be valuable, even if this coefficient applies only to the Bell System and to other companies having approximately the same employment practices.

An article[1] by Mr. Walter S. Gifford, President of the Bell System, contains a discussion of some of the relationships between scholarship and salaries. President Gifford, however, did not determine in the case of the Bell System a coefficient of correlation between scholarship and salaries.

The purpose of this article is not a new contribution to statistical method, but is an application of the method[2] of finding the coefficient of correlation when the two variables have not been quantitatively measured. This method will be applied to the chart on page 672 of President Gifford's article, in order to determine for the Bell System the coefficient of correlation between scholarship and salaries.

FINDING THE COEFFICIENT OF CORRELATION

An explanation of the chart. It is based on the experience of 2,144 Bell System employees over five years out of college. First, assume these employees

[1] It is entitled "Does Business Want Scholars?" and was printed in the May 1928 issue of Harper's Magazine.

[2] It can be found in Elderton's "Frequency Curves and Correlation."

are grouped according to their grades in college. In the high scholarship group put those who graduated in the highest third of their classes. The middle and low scholarship groups are formed in like manner. Secondly, suppose the same employees are divided into three equal groups according to their salaries. Then, the salary of any one of the employees would be high, middle, or low.

Assume a hypothetical group of 300 employees who are college graduates. Suppose that the scholarship of 100 of them was high, that the scholarship of 100 of them was middle, and that the scholarship of the others was low. Also assume that the salary experience of these 300 employees is the same as that of the 2,144 employees of the Bell System.

The 300 employees can be grouped according to the following table.

TABLE NO. 1

Salary	Scholarship			Totals
	Low	Middle	High	
High	22	24	48	94
Middle	31	39	27	97
Low	47	37	25	109
Totals	100	100	100	300

This table can be combined as follows.

TABLE NO. 2

Salary	Scholarship	
	Low & Middle	High
High	c	d
Middle & Low	a	b

Then, $c = 46$, $a = 154$, $d = 48$, and $b = 52$. Assume $N = 300$.

Assume x is a function of grades received in college. Suppose y is a function of salaries received. Assume that the frequencies x and y both follow the normal curve of error whose standard deviation is equal to one. Also assume that the average of x and the average of y are both equal to zero. It is a matter of common knowledge that salaries are not arranged in a symmetrical fashion; y is not a linear function of salaries.

In the formulas which follow, r is the symbol for the coefficient of correlation. These formulas are applied to Table No. 2. We have

$$\frac{1}{\sqrt{2\pi}}\int_0^h e^{-\frac{1}{2}x^2}\,dx = \frac{(a+c)-(b+d)}{2N} = .167, \text{ and } h = .4316.$$

Also

$$\frac{1}{\sqrt{2\pi}}\int_0^k e^{-\frac{1}{2}y^2}\,dy = \frac{(a+b)-(c+d)}{2N} = .187, \text{ and } k = .4874.$$

Then,

$$H = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}h^2} = .3635, \text{ and } K = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}k^2} = .3543.$$

All the quantities except r in the following approximate equation are known:

$$\frac{ad-bc}{N^2HK} = r + \frac{r^2}{2}hk + \frac{r^3}{6}(h^2-1)(k^2-1)$$
$$+ \frac{r^4}{24}h(h^2-3)k(k^2-3) + \frac{r^5}{125}(h^4-6h^2+3)(k^4-6k^2+3).$$

Therefore,

$$.0261r^5 + .0681r^4 + .1034r^3 + .1052r^2 + r - .4314 = 0.$$

Then, r is approximately equal to .4051. Consequently, for practical purposes we can assume that $r = .4$.

28 Boody Street
Brunswick, Maine

John L. Roberts

NOTE ON THE DERIVATION OF THE MULTIPLE CORRELATION COEFFICIENT

Consider N observed values of each of n variables. These $n \cdot N$ values may be tabulated in a double-entry table as follows:

$$\begin{array}{c} X_{11}\ X_{12}\ X_{13}\cdots X_{1N} \\ X_{21}\ X_{22}\ X_{23}\cdots X_{2N} \\ \\ \\ X_{n1}\ X_{n2}\ X_{n3}\cdots X_{nN} \end{array}$$

where X_{ik} is the k^{th} value of the i^{th} variable.

Using the i^{th} variable as the dependent variable, the general linear relationship between the n variables may be expressed by

$$x_i = {}_ia_1\,x_1 + {}_ia_2\,x_2 + \cdots + {}_ia_{i-1}\,x_{i-1} + {}_ia_{i+1}\,x_{i+1} + \cdots + {}_ia_n\,x_n \qquad (1)$$

where

${}_ia_j$ is the general parameter which is to be determined empirically;

$x_j = X_j - M_j$;

M_j is the arithmetic mean of the j^{th} variable.

By the method of least squares, the constants of (1) must satisfy the normal equations:

$$(\Sigma x_1^2)_i a_1 + (\Sigma x_1 x_2)_i a_2 + \cdots + (\Sigma x_1 x_{i-1})_i a_{i-1}$$
$$+ (\Sigma x_1 x_{i+1})_i a_{i+1} + \cdots + (\Sigma x_1 x_n)_i a_n = \Sigma x_1 x_i$$

$$(\Sigma x_2 x_1)_i a_1 + (\Sigma x_2^2)_i a_2 + \cdots + (\Sigma x_2 x_{i-1})_i a_{i-1}$$
$$+ (\Sigma x_2 x_{i+1})_i a_{i+1} + \cdots + (\Sigma x_2 x_n)_i a_n = \Sigma x_2 x_i$$

. . . .

. . . .

. . . .

$$(\Sigma x_{i-1} x_1)_i a_1 + (\Sigma x_{i-1} x_2)_i a_2 + \cdots + (\Sigma x_{i-1} x_n)_i a_n = \Sigma x_{i-1} x_i$$

$$(\Sigma x_{i+1} x_1)_i a_1 + (\Sigma x_{i+1} x_2)_i a_2 + \cdots + (\Sigma x_{i+1} x_n)_i a_n = \Sigma x_{i+1} x_i$$

. . . .

. . . .

$$(\Sigma x_n x_1)_i a_1 + (\Sigma x_n x_2)_i a_2 + \cdots + (\Sigma x_n^2)_i a_n = \Sigma x_n x_i$$

where

$$\left(\sum x_i x_j\right) = \sum_{k=1}^{N} (X_{ik} - M_i)(X_{jk} - M_j).$$

But

$$(\Sigma x_i x_j) = N r_{ij} \sigma_i \sigma_j,$$
$$(\Sigma x_i^2) = N \sigma_i^2 = N r_{ii} \sigma_i \sigma_i \qquad (2)$$

where

r_{ij} is the Pearsonian coefficient of correlation between the i^{th} and j^{th} variables,

σ_i, the standard deviation of the i^{th} variable.

Substituting the right members of (2) in the normal equations, we obtain the system:

$$\sum_{k=1}^{n} r_{1k} \sigma_1 \sigma_k \, {}_i a_k = 0$$

$$\sum r_{2k} \sigma_2 \sigma_k \, {}_i a_k = 0$$

$$\sum_{k=1}^{n} r_{i-1,k} \, \sigma_{i-1} \sigma_k \, {}_i a_k = 0 \qquad (3)$$

$$\sum_{k=1}^{n} r_{i+1,k} \, \sigma_{i+1} \sigma_k \, {}_i a_k = 0$$

$$\sum_{k=1}^{n} r_{nk} \sigma_{1,} \sigma \, {}_i a_k = 0$$

where

$$ {}_i a_i = -1. $$

Let

$$ A = \begin{vmatrix} r_{11}\sigma_1\sigma_1 & \cdots & r_{n1}\sigma_n\sigma_1 \\ \\ r_{1n}\sigma_1\sigma_n & & r_{nn}\sigma_n\sigma_n \end{vmatrix} \tag{4} $$

A_{ij} be the first minor of the element $r_{ij}\sigma_i\sigma_j$ in A, ${}_{ik}A$ be A with the i^{th} and k^{th} columns interchanged, and ${}_{ik}A_{ii}$ be the first minor of the element in the i^{th} column and i^{th} row of ${}_{ik}A$.

Solving (3) for ${}_i a_k$ by Cramer's rule, we find

$$ {}_i a_k = \frac{{}_{ik}A_{ii}}{A_{ii}}. $$

But it can easily be proved that

$$ {}_{ik}A_{ii} = (-1)^{i-k+1} A_{ik}; $$

hence

$$ {}_i a_k = (-1)^{i-k+1} \frac{A_{ik}}{A_{ii}}. $$

Using cofactors of A instead of minors, we have

$$ {}_i a_k = (-1)^{i-k+1} \frac{(-1)^{i+k} D_{ik}}{D_{ii}} = -\frac{D_{ik}}{D_{ii}}. $$

Without writing the determinant out in full, we notice that the σ's can be factored out. Hence

$$ \begin{aligned} {}_i a_k &= -\frac{\sigma_1^2\sigma_2^2 \cdots \sigma_{k-1}^2\sigma_k\sigma_{k+1}^2 \cdots \sigma_{i-1}^2\sigma_i\sigma_{i+1}^2 \cdots \sigma_n^2 K_{ik}}{\sigma_1^2\sigma_2^2 \cdots \sigma_{i-1}^2\sigma_{i+1}^2 \cdots \sigma_n^2 K_{ii}} \\ &= -\frac{\sigma_i K_{ik}}{\sigma_k K_{ii}}, \end{aligned} \tag{5} $$

where

$$ K = \begin{vmatrix} r_{11} & \cdots & r_{1n} \\ \\ r_{n1} & & r_{nn} \end{vmatrix} $$

Using these derived values for the coefficients, we may write (1) in the symmetric form:

$$\frac{K_{i1}}{\sigma_1}(X_1 - M_1) + \frac{K_{i2}}{\sigma_2}(X_2 - M_2) + \cdots + \frac{K_{in}}{\sigma_n}(X_n - M_n) = 0,$$

or

$$\sum_{j=1}^{n} \frac{K_{ij}x_j}{\sigma_j} = 0. \tag{6}$$

For a multiple correlation coefficient, we use the formula

$$R_i^2 = 1 - \frac{\sum_{j=1}^{N}\left[x_{ij} - \left(\sum_{k=1}^{i-1} {}_i a_k x_{kj} + \sum_{k=i+1}^{n} {}_i a_k x_{kj}\right)\right]^2}{N\,\sigma_i^2}$$

which measures the amount of observed dispersion from the regression plane in which X_i is the dependent variable.

Substituting the values for the a's, we find

$$R_i^2 = 1 - \frac{\sum_{j=1}^{N}\left(\frac{K_{i1}x_{1j}}{\sigma_1} + \frac{K_{i2}x_{2j}}{\sigma_2} + \cdots + \frac{K_{in}x_{nj}}{\sigma_n}\right)^2}{K_{ii}^2\,N}.$$

Squaring the bracket expression and using (2), we obtain

$$R_i^2 = 1 - \frac{1}{K_{ii}^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\left(\frac{K_{ik}K_{il}\sum x_{kj}x_{lj}}{N\sigma_k\sigma_l}\right)$$

$$= 1 - \frac{1}{K_{ii}^2}\left[\sum_{k=1}^{n}\sum_{l=1}^{n} K_{ik}K_{il}r_{kl}\right]$$

$$= 1 - \frac{1}{K_{ii}^2}\left[\sum_{k=1}^{n}\left(K_{ik}\sum_{l=1}^{n} K_{il}r_{kl}\right)\right].$$

The second sum is the sum of the products of the elements in the k^{th} row by the cofactors of the elements in the i^{th} row. This sum is necessarily zero unless $k = i$; but if $k = i$, this sum is equal to K.

$$R_i^2 = 1 - \frac{1}{K_{ii}^2}(K_{ii}\,K) = 1 - \frac{K}{K_{ii}}.$$

Oregon State Agriculture College
School of Science
Corvallis, Oregon

William J. Kirkham

NOTE ON NUMERICAL EVALUATION OF DOUBLE SERIES[1]

1. The Euler-Maclaurin summation formula has been extended to two variables by Dr. Sheppard,[2] and Mr. Irwin,[3] to determine cubature formulas. A more complicated two-dimensional form was given by Baten[4] involving product polynomials, for which a remainder term was also calculated. The purpose of this note is to apply the simpler formula to the numerical evaluation of double series of positive terms. The method may be extended to multiple series of order $p > 2$. If the double series converges one may sum by rows (or columns), using the ordinary sum formula twice. The method is to take out a rectangular block of mn terms and then apply the formula to the remaining terms. By taking m and n sufficiently large one may cause the series resulting from the formula to converge sufficiently rapidly to obtain the sum to the desired number of decimal places. For practical work the error may be estimated because of the asymptotic character of the series involved in the Euler-Maclaurin formula.

Write this in the form

$$
\begin{aligned}
\sum_{a}^{s-1} f(x) &= \int_a^s f(x)dx + \tfrac{1}{2}f(a) - \tfrac{1}{2}f(s) - \frac{f'(a)-f'(s)}{12} + \frac{f'''(a)-f'''(s)}{720} - \\
&\frac{f^{\mathrm{V}}(a)-f^{\mathrm{V}}(s)}{30240} + \frac{f^{\mathrm{VII}}(a)-f^{\mathrm{VII}}(s)}{1209600} - \cdots + (-1)^r B_r \frac{f^{(2r-1)}(a)-f^{(2r-1)}(s)}{(2r)!} + \cdots .
\end{aligned}
\tag{1}
$$

If $s \to \infty$ one has accordingly in the ordinary case of convergence

$$
\sum_{a}^{\infty} f(x) = \int_a^\infty f(x)dx + \tfrac{1}{2}f(a) - \frac{f'(a)}{12} + \frac{f'''(a)}{720} - \frac{f^{\mathrm{V}}(a)}{30240} + \cdots . \tag{2}
$$

Now define $v(x) = \sum_{y=b}^{\infty} u(x, y) = \int_b^\infty u(x, y)\,dy + \tfrac{1}{2}u(x, b) - \frac{u_y(x, b)}{12} + \frac{u_{y^3}(x, b)}{720} - \cdots$ and $w(y) = \sum_{x=a}^{\infty} u(x, y) = \int_a^\infty u(x, y)dx + \tfrac{1}{2}u(a, y) - \frac{u_x(a, y)}{12} + \frac{u_{x^3}(a, y)}{720} - \cdots$, then $\sum_{x=1}^{\infty}\sum_{y=1}^{\infty} u(x, y) = \sum_{x=1}^{a-1}\sum_{y=1}^{b-1} u(x, y) + \sum_{x=1}^{\infty} v(x) + \sum_{y=1}^{b-1} w(y)$

$$
\begin{aligned}
&= \int_1^\infty v(x)dx + \tfrac{1}{2}v(1) - \frac{v'(1)}{12} + \frac{v'''(1)}{720} - \frac{v^{\mathrm{V}}(1)}{30240} + \cdots + \sum_{x=1}^{a-1}\sum_{y=1}^{b-1} u(x, y) \\
&+ \int_1^b w(y)dy - \tfrac{1}{2}w(b) + \tfrac{1}{2}w(1) + \frac{w'(b)-w'(1)}{12} - \frac{w'''(b)-w'''(1)}{720} +
\end{aligned}
\tag{3}
$$

[1] Presented to the Society, Nov. 30, 1934.

[2] W. F. Sheppard, "Some Quadrature Formulae," Proc. London Math. Soc., Vol. xxxii, 1900.

[3] J. O. Irwin, "Tracts for Computers," No. X, Cambridge Univ. Press, 1923, On Quadrature and Cubature.

[4] W. D. Baten, "A Remainder for the Euler-Maclaurin summation formula in two independent variables," Amer. Journal of Math., Vol. 54, 1932, pp. 265–275.

Instead of this one may use $\sum_{x=1}^{a-1} \sum_{y=1}^{b-1} u(x, y) + \sum_{y=1}^{\infty} w(y) + \sum_{x=1}^{a-1} v(x)$. The scheme of the double series may be illustrated by a sketch of a quadrant of the xy-plane in which the point (x, y) represents the term $u(x, y)$.

Evidently by taking a combination of results from (3) one may evaluate quite readily such finite sums as $\sum_{x=p}^{q} \sum_{y=r}^{t} u(x, y)$ where q and t are large.

As an illustration of (3) consider $\sum \sum (x^2 + a^2y^2)^{-2}$. Here one needs to evaluate the integral of the summand. The transformations $x = ay \tan \theta$ and $y = 1/t$ lead to a form which may be integrated by parts. The more complicated form $\sum \sum (ax^2 + 2bxy + cy^2)^{-s}$ for the case in which $s > 3/2$, $a > 1$, might be handled by using $x = 1/t$ and approximate integration by Simpson's rule.

Take as a second example $\sum \sum (x + y)^{-p}$, $p > 2$. The case of $p = 4$ was carried out by taking $a = b = 10$ in (3) and carrying the computation to twelve decimals. The series involved converge rapidly and a result was obtained which differed by 2 in the 12th place from the true value $0.119\ 733\ 669\ 448^{+}$. By summing diagonally one may convert this to the simple series $\sum_{1}^{\infty} z(z + 1)^{-4}$ or $\sum_{2}^{\infty} (s - 1)s^{-4} = \sum_{1}^{\infty} (s^{-3} - s^{-4})$. The method of summation diagonally may be extended to $\sum \sum (x + ay)^{-p}$, $p > 2$, $a > 0$, by the applications of the Euler-Maclaurin sum formulas (1), (2) in succession after a triangular array of terms have been omitted.

The form $\sum \sum x^{-p} y^{-q}$ can be written as the product of the single series $(\sum x^{-p})(\sum y^{-q})$.

2. Another method of numerical evaluation is the analog of that used for single series by the author.[5] Instead of rectangles one has right prisms of square or rectangular cross-section. Instead of shifting the rectangles one unit to the right to determine upper and lower bounds the prisms are shifted diagonally so that they go effectively one unit in each variable. In the case of a square base each prism is moved along the 45° line one diagonal unit length. For the lower bound instead of trapezoids one uses truncated prisms. For example, the prism of height u_{mn} is cut by two planes, one determined by the upper vertices u_{mn}, $u_{m,n+1}$, $u_{m+1,n}$ and the other by the upper vertices $u_{m+1,n}$, $u_{m,n+1}$, $u_{m+1,n+1}$ of the truncated prism. The surface $z = u(m, n)$ passes through all the upper corners of the truncated prisms. Each prism is composed of two truncated triangular prisms. Now the volume of such a triangular prism is the arithmetic mean of its vertical edges multiplied by the area of its base.

[5] "A New Method for Finding the Numerical Sum of an Infinite Series," Amer. Math. Monthly, vol. XL, No. 9, Nov., 1933, pp. 537–542.

Hence the difference in volume between the truncated rectangular prism mentioned above and the prism of uniform hight $z = u_{mn}$ can be shown to be

$$(4)\qquad (5u_{mn} - u_{m+1,\,n+1} - 2u_{m+1,\,n} - 2u_{m,\,n+1})/6.$$

Let us consider series whose corresponding surfaces do not rise above these truncated prisms. This sort of truncated prism differs less from the volume under the surface than the one formed by the diagonal joining the other pair of upper vertices and planes through it for upper faces. The lower bound for the remainder is the volume under the surface extending to infinity in the m and n directions plus the sum of these differences. Accordingly one determines as the lower bound for the remainder $R_{m-1,\,n-1}$ after summing a rectangular array $\sum_{i=1}^{m-1}\sum_{j=1}^{n-1} u_{i,\,j}$ the form

$$(5)\qquad \begin{aligned}&(2u_{m,1} + 2u_{1,n} - 5u_{m,n})/6 + \tfrac{1}{2}\sum_{i=1}^{\infty} u_{m+i,1} + \tfrac{1}{2}\sum_{j=1}^{\infty} u_{1,n+j} + \tfrac{1}{2}\sum_{i=1}^{m} u_{i,n}\\ &\quad + \tfrac{1}{2}\sum_{j=1}^{n} u_{m,j} + \int_1^{\infty}\int_m^{\infty} u_{m,n}\,dmdn + \int_n^{\infty}\int_1^{m} u_{m,n}\,dmdn < R_{m-1,n-1}.\end{aligned}$$

The upper bound may likewise be given as follows:

$$(6)\qquad R_{m-1,n-1} < S + T + \int_{n-1}^{\infty}\int_{m-1}^{\infty} u_{m,n}\,dmdn - k\left[\sum_{j=n-1}^{\infty} u_{m-1,j} + \sum_{i=m}^{\infty} u_{i,n-1}\right]$$

where

$$(7)\qquad S = \sum_{i=m}^{\infty}\sum_{j=1}^{n-1} u_{ij}, \qquad T = \sum_{j=n}^{\infty}\sum_{i=1}^{m-1} u_{ij},$$

$$(8)\qquad k = \frac{\int_{n-1}^{\infty}\int_{m-1}^{\infty} u_{m,n}\,dmdn - \int_{n}^{\infty}\int_{m}^{\infty} u_{m,n}\,dmdn - \sum_{j=n-1}^{\infty} u_{m-1,j} - \sum_{i=m}^{\infty} u_{i,m-1}}{u_{m-1,n-1} + u_{m,n-1}}.$$

An alternate definition of k is

$$(9)\qquad k' = \left[\int_{n-1}^{n}\int_{m-1}^{m} u_{m,n}\,dmdn - u_{m,n}\right] \div (u_{m-1,n-1} - u_{m,n}).$$

An illustration is afforded by $\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} (m+1)^{-4}$ for which $k = .45614$, $k' = .44536$ when $m = n = 10$ in (8), (9). In this case (5) gave an error of -14×10^{-6} and (6) an error of 10^{-6}.

S and T may be evaluated by the method published in the Monthly.[6]

One must assume that k increases with m and n. It is evident that for this

[6] Loc. cit.

method and for the one in the Monthly differentiability is not required but only integrability, conditions less restrictive than those required by the Euler-Maclaurin summation formulas. It is also clear that the method may be extended to multiple series of positive terms of multiplicity greater than two.

Department of Mathematics
University of Nebraska
Lincoln, Nebraska

Chester C. Camp

REPORT OF THE ANNUAL MEETING OF THE INSTITUTE OF MATHEMATICAL STATISTICS

The meeting of the Institute of Mathematical Statistics for 1936 was held in Chicago on December 28–30 in connection with the meetings of the American Statistical Association and the Econometric Society.

In addition to the sessions at which voluntary papers were read, a session with invited papers was held on the morning of December 30. At the invitation of the Program Committee, Professor P. R. Rider presented a paper on "Recent Advances in Mathematical Statistics: Factorial Design" and Professor Harold Hotelling spoke on "The Analysis of Sets of Correlated Variates."

Professor C. C. Craig of the University of Michigan and Professor A. R. Crathorne of the University of Illinois constituted the Program Committee.

At the business meeting of the Institute, the following officers were elected for the year 1937: President, Dr. W. A. Shewhart; Vice-Presidents, Professors P. R. Rider and B. H. Camp; Secretary-Treasurer, Professor A. T. Craig.

The Institute voted that it would presumably hold its 1937 meeting with the American Mathematical Society.

ALLEN T. CRAIG,
Secretary.

NOTICE TO SUBSCRIBERS

Plans are under way to include in the ANNALS a new section, entitled "Numerical Illustrations of Statistical Methodology." This new section will be a regular feature of the ANNALS, and will deal with the application of statistical technique and theory to the solution of problems in various fields. It is hoped that this new section will be of considerable value to those who are primarily interested in numerical applications of the more recent theoretical developments in mathematical statistics.

The Editor will welcome contributions to this new section of the ANNALS.

REGRESSION AND CORRELATION EVALUATED BY A METHOD OF PARTIAL SUMS

By Felix Bernstein

"To be sure, Laplace viewed the matter in a similar way but he selected the absolute value of the error as a measure of loss. But if we mistake not, this position is certainly not less arbitrary than our own; that is to say, whether the double error is to be considered just as tolerable as, or worse than, the simple error twice repeated and whether it is thus more fitting to ascribe to the double error only a double weight, or a greater one, is a question which is neither in itself clear nor determinable by mathematical proof but has to be left entirely to individual discretion.

"Furthermore, it cannot be denied that the assumption under discussion violates the principle of continuity and precisely for this reason the procedure based on it strongly defies analytic treatment while the results to which our principle leads have the advantage of simplicity as well as of generality."—
F. G. Gauss: Theoria combinationis observationum, pars prior, art. 6.

Since the "Theoria Combinationis" of C. F. Gauss appeared in the year 1821 a century of Mathematical Statistics has been dominated by the ideas of this classical treatise—ideas whose fertility does not seem to be exhausted even today.

The germ of most modern contributions to mathematical statistics—in fact also those of Karl Pearson and his school—go back decidedly to this paper. Though the immediate achievements of Gauss are so conspicuous as not to need any comment, a true critical appreciation of the work can be gained only by comparing it with the previous methods of Laplace, superseded by those of Gauss.

For such critical appreciation, C. F. Gauss himself has prepared the ground in the lines quoted at the beginning of this article. To Gauss the standard deviation is a measure of uncertainty or risk of a game in which the errors of observation are considered as causing only losses. In this he follows the lead of his great predecessor. The difference between them is that Gauss adopts the square of the error as a measure of the loss while Laplace adopts its absolute value for this purpose. Either choice frees the error from its sign so that the loss is the same regardless of the sign of the error.

Gauss considers this choice of the measure of the loss as purely conventional. Therefore he feels justified in adopting the square of the error because in adopting the square instead of the absolute value of the error, the mathematics he uses remains in the easily accessible domain of analytical processes. This creates for these methods a superiority in elegance, simplicity, and generality.

The modern developments of mathematical statistics, based on the principles

of Gauss, have confirmed the correctness of this viewpoint. This has proved true particularly in the theory of analysis of variance developed by R. A. Fisher and in the more general theory of semi-invariants, first defined by N. H. Thiele.

The inadequacy of the Gaussian method seriously impairing its value for statistical use has come to light through the investigations of Karl Pearson of distributions of one and two variables. Since the moments of higher order involve standard deviations of increasing magnitude the characterization of the distributions by means of the moments, in line with the Gauss-Thiele concepts, becomes practically impossible. Therefore it was of the greatest interest that Lindeberg was able to derive an expression for the standard deviation of a measure of skewness constructed not on Gaussian but on Laplacian lines, namely based exclusively upon the sign of the error. The mathematical difficulties surmounted by Lindeberg by a very involved and difficult analysis—with some clearly indicated gaps in the proofs—are precisely of the character of those that Gauss wished to avoid. Encouraged by the success of Lindeberg, I have developed in two papers[1] the standard deviations of more general moments and the correlations between them of which the mean deviation of Laplace and Lindeberg's measure of skewness are special cases. The proofs have been arrived at by a rather simple and rigorous procedure. These new moments, together with the old ones, form a new system of statistical characteristics by which a distribution in one or two variables can be described by expressions of lower order and therefore of greater precision. This method makes unnecessary the use of moments of higher order than the third.

But another point of interest is still involved. It has been assumed that the Gaussian characteristics give a greater amount of information than those of Laplace. This is proved, however, only for the case of the normal distribution $\frac{h}{\sqrt{\pi}} e^{-h^2x^2}$. This was recognized by Gauss himself in his paper of April, 1816, that appeared five years earlier than the Theoria Combinationis Observationum. In article 6 of his paper, he says, that the constant h of a normal distribution obtained from one hundred observations by the use of the standard error is as exact as that obtained from one hundred fourteen observations in which the mean deviation is used. Hence with a given number of observations only the equivalent of 88% of the total are used by the second method. This does not hold true for all distributions. The following theorem can easily be proved: The amount of information as defined above, furnished by the use of the mean deviation is greater, equal to, or less than that furnished by the standard deviation, depending respectively upon whether

[1] Felix Bernstein: "Die mittleren Fehlerquadrate und Korrelationen der Potenzmomente und ihre Anwendung auf Funktionen der Potenzmomente," Metron, Vol. X, N. 3 (Nov. 1932).

Felix Bernstein: "Uber den mittleren Fehler der Potenzmomente." Zeitschr. f. d. ges. Vers.-Wissenschaft, Bard 30, Heft 3, March 1930.

$$(\beta_2 - 1) \gtreqless 4(\beta_0 - 1)$$

where

$$\beta_0 = \frac{\mu_2}{\vartheta^2}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

μ_k the k-th moment and ϑ = the mean deviation.

For example, in the distribution $\frac{h}{2} e^{-h|x|}$, the mean deviaition furnishes a greater amount of information than the standard deviation.[2]

In the present paper, we shall discuss the practical use of expressions for correlation and regression in which the new type of statistics formed along Laplacian lines will be used. These new expressions are of a linear form and can be computed therefore more easily than those of Karl Pearson. The amount of information given by these expressions is less than that given by the expressions of Pearson if the normal law, in two variables, is fulfilled. For other distributions, however, this is not generally true. The determination of the standard deviations of these new expressions is given in Metron.[3]

The application of the new expressions of regression and correlation to grouped data is set forth here for the first time. The method is strongly recommended for all cases in which the data lose reliability with increasing deviations from the mean. Deviations in the new method enter the expressions only in the first degree and not in the second as in the case of Pearson's. It is obvious that the influence of the doubtful extreme readings is, therefore, considerably lessened. Since our expressions are linear, no adjustments for grouping (Sheppard's corrections) are necessary.

It ought to be mentioned here that linear expressions for the measurement of correlation have been set up before.

K. Pearson (Biometrika) and Egon Pearson (Biometrika) have derived an expression called "linear correlation ratio" which in case of linear regression is identical with the correlation coefficient.

K. Pearson also discusses the linear correlation coefficient

$$r = \frac{1}{2}\left(S\frac{ysgx}{xsgx} + S\frac{xsgy}{ysgy}\right),$$

[2] To this second type of distribution curves also belongs $y = \psi(x)$ where $x(x)$ is the mean of two Gaussian curves with the same origin, i.e. $\psi(x) = \frac{1}{2}\left(\frac{h}{\sqrt{\pi}} e^{-h^2x^2} + \frac{kh}{\sqrt{\pi}} e^{-h^2k^2x^2}\right)$ $1.6 < k < 3.4$.

I owe this remark and some other valuable suggestions regarding the subject of this paper to Mr. Myron Fuchs.

[3] *Op. cit.*

suggested by Lenz and various other linear expressions, all similar to our expression (1). He finds that they are all equal to his quadratic correlation coefficient in the case of a Gaussian distribution.

However, their expressions were not recommended by those authors for the determination of correlation between quantitative variables, because—

1. No easy and practicable methods were given for their evaluation in the case of grouped data.
2. Their standard deviations were not determined.

We now proceed to define the new formulas and to describe the methods for their evaluation. The proofs are furnished in the Appendix to this paper.

Let r_1 and r_2 denote the regression coefficients of x on y and y on x respectively, and r, as usual, the coefficient of correlation, and by $\bar{x}$ and $\bar{y}$ the arithmetic means of the x's and y's. Let us take $\bar{x}$, $\bar{y}$ as the origin, so that x, y are the deviations from the mean. We have

$$(1) \qquad \begin{aligned} r_1 &= \frac{\underset{+y}{S}x}{\underset{+y}{S}y} \quad \text{or} \quad r_1 = \frac{\underset{-y}{S}x}{\underset{-y}{S}y} \\ r_2 & \quad \frac{\underset{+x}{S}y}{\underset{+x}{S}x} \quad \text{or} \quad r_2 \quad \frac{\underset{-x}{S}y}{\underset{-x}{S}x} \\ r &= \end{aligned}$$

$\underset{+y}{S}x$ denotes a partial sum of the x's, this sum being extended over all the x's of the observations whose y is positive and the other sums have a corresponding meaning.

It should be noted though that if data occur whose y-deviation is 0 (practically never in a grouped table) one-half of the sum of these x's should be added to $\underset{+y}{S}x$. In the $\underset{+x}{S}$ a similar addition should be made in case observations occur in which x is zero. (See Table IV.)

The formulas (1) and all following ones will be proved in the appendix to this article.[4]

[4] Using r_1 and r_2 of (1) the regression lines are $y = r_2x$ and $x = r_1y$. They are those straight lines which fit the data best according to the method of least squares, if the weight of the deviations is taken inversely proportional to the absolute value of the variable. Taking x for instance as the independent variable, r_2 is the value of m which minimizes $S\frac{1}{|x|}(y - mx)^2$ (the sum extended over all data x y).

The standard deviations of r_1 and r_2 are

(2)
$$\sigma_{r_1}^2 = \frac{r_1^{2r}}{2N}(1 + m(m - 2r)) \qquad \text{where } m = \frac{\underset{+x}{Sx}}{\underset{+y}{Sx}}$$

$$\sigma_{r_2}^2 = \frac{r_2^{2r}}{2N}(1 + n(n - 2r)) \qquad \text{where } n = \frac{\underset{+y}{Sy}}{\underset{+x}{Sy}}$$

We are now going to illustrate the computation of r and for this purpose we shall use a table of Pearson's which gives the correlation between the heights of fathers and daughters.

The totals at the right and lower end of the table are first computed and the bracketed numbers are the sums of the numbers that precede. The means are

$$\bar{x} = \frac{1659.5 - 1179}{1376} = +\frac{480.5}{1376}$$

and

$$\bar{y} = \frac{1650.9 - 1390}{1376} = +\frac{260.5}{1376}$$

whose signs determine on which side of the working mean to "quarter" the table. This quartering is done in Table 1 by the lines vv and hh. Then the totals above the heavy horizontal separating line hh and those to the left of the vertical separating line vv are found, e.g. 2, 4.5, 7.25, $\cdots$ and .5, .5, 0, $\cdots$. Multiplying these totals by the respective class marks, we find the outside lines: 18, 36, 50.75, $\cdots$ and 5.5, 5, 0, $\cdots$.

$\underset{-y}{Sx}$ is now $= 1107.5 - 420.5 = 687$, and an adjustment for the fact that a working mean has been used has yet to be made. This adjustment is $\bar{x}N_{-y}$ where N_{-y} is the number of negative y's. ($N_{-y} = 728$.)

We have therefore for the adjusted values

$$\underset{-y}{Sx} = 1107.5 - 420.5 + \frac{260.5}{1376}\cdot 728 = 825.07$$

$$\underset{-y}{Sy} = 1179 + \frac{480.5}{1376}\cdot 728 = 1433.21$$

$$r_1 = .5757 \qquad\qquad r_2 = .5170$$

$$r = .546$$

The standard deviations, according to the formulas (2) are

$$\sigma_{r_1} = .031 \qquad \sigma_{r_2} = .027$$

TABLE 1

Correlation between Heights of Fathers and Daughters

x → Height of Fathers y ↓ Height of Daughters

In Inches

v

			18	36	50.75	66	206.25	183	202.5	212.5	132.5	(1107.5)	86.25	116.5	115.5	69	23.75	7.5	1.75		(420.25)	
	Totals above line		2	4.5	7.25	11	41.25	45.75	67.5	106.25	132.5		86.25	58.25	38.5	17.25	4.75	1.25	.25		Totals	
	Totals left of line		−9	−8	−7	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	(728)	
5.5	.5	−11					.25	.25													.5	5.5
5	.5	−10					.25	.25													.5	5
		−9																				
8	1	−8								1											1	8
31.5	4.5	−7	.25	.25		.25	1.25	.5		1	.5	.5									4.5	31.5
81	13.5	−6	.25	.25	.5	1.5	4.5	1	1.5	1.5	2.5		.5	.5							14.5	88
73.75	14.75	−5	.25	.75	.5	.75	.75	1	1.75	1.25	5	2.75	.5	.25							15.5	77.5
163	40.75	−4	.5	1	2		6	4.75	5	6.25	11.75	3.5	3.5	2	1.75	.5					48.5	194
240.75	80.25	−3	.75	.75		2.5	8	6.25	12.5	18.25	20.25	11	9	4.75	2.5	1.25	1.25				99	297
212.5	106.25	−2		.5	1.75	2	9.75	11.5	13	23.75	23.75	20.25	16.5	10.25	4.25	3	1.25				141.5	283
131.75	131.75	−1		1	2.25	2	4.5	12	22.75	26	33	28.25	24.75	14.25	13.75	4.75	.75	.5			190.5	190.5
(952.75)		0			.25	2	6	8.25	11	22.75	35.75	37.25	31.5	26.25	16.25	7.75	1.5	.75	.25		212	(1179)
87.25	87.25	1			.25	2.5	1.75	3.25	9.25	23	18.75	28.5	33	34.25	24.5	11.75	5.5	1	.25	1	198.5	198.5
108.5	54.25	2				.5	1	.5	11	12.25	9.25	19.75	30	26.5	22.25	15	4.75	3.75	2	1	150.5	319
113.25	37.75	3				.5	.5	1.5	3.25	7.25	8.75	16	26.25	26.75	20.5	18.5	7.75	4.25	.25	.5	142.5	427.5
71	17.75	4							1	5.75	7	4	14.25	13.25	12	11.25	4.5	3.75	.75		77.5	310
22.5	5.5	5					.25	.25	.25	.25	1.5	3	5.5	4.25	5.75	5.25	3.75	2.5	1.5	2	36	180
9	1.5	6					.25	.25	.25	.25	.25	.25	1	2.5	6.5	2.25	2.75	2	1		19.5	117
		7											1.75	.25	4.5	.75	1.25	.75	.25		9.5	66.5
		8											.5		.5	.5	1.5	.75	.25		4	32
		9											1								1	9
(416.5)	Totals	(728.5)	2	4.5	7.5	14.5	45	51.5	92.5	155	178	175	199.5	166	135	82.5	36.5	20	6.5	4.5	(1376)	(1650.5)
			18	36	52.5	87	225	206	277.5	310	178	(1390)	199.5	332	405	330	182.5	120	45.5	36	(1650.5)	

h

v

Working Mean x = 67.5 y = 63.5

Class width 1 Inch

The standard deviation of $r^2 = r_1 \times r_2$ has to be estimated by using the general formula for the standard deviation of the product c of two variables a and b;

$$\frac{\sigma_c^2}{c^2} = \frac{\sigma_a^2}{a^2} + \frac{\sigma_b^2}{b^2} + \frac{2R\sigma_a\sigma_b}{ab}$$

R being the correlation coefficient between a and b. Since $-1 < R < +1$, substitution of these limits for R leads to the inequalities

$$\left(\frac{\sigma_a}{a} - \frac{\sigma_b}{b}\right)^2 < \frac{\sigma_c^2}{c} < \left(\frac{\sigma_a}{a} + \frac{\sigma_b}{b}\right)^2$$

putting $a = r_1$, $b = r_2$, $c = r^2$ we have

$$\frac{\sigma_{r_1}}{r_1} - \frac{\sigma_{r_2}}{r_2} < \frac{\sigma_{r^2}}{r} < \frac{\sigma_{r_1}}{r_1} + \frac{\sigma_{r_2}}{r_2}$$

Considering the relation $\sigma_r = \dfrac{\sigma_{r_2}}{2r}$

we have $2r(\sigma_{r_1} r_2 - \sigma_{r_2} r_1) < \sigma_r < 2r(\sigma_{r_1} r_2 + \sigma_{r_2} r_1)$
from which we derive with sufficient approximation

$$\sigma_r < \cdot 030$$

A slightly different arrangement for computing r has been made in the following table.

TABLE II

*Correlation between diameter of the stem and length of the lonest flower petal of Trientalis europaea**

	PS	3	15	34	45	30	6	2	0	0	0	0	
PS		−4	−3	−2	−1	0	1	2	3	4	5	6	Total
1	−4	1											1
7	−3	1	4	1	1								7
29	−2	1	9	16	3	1							30
33	−1		2	9	22	9	2	1					45
27	0			8	19	20	4	1					52
8	1	1			7	18	12	6	4				48
1	2				1	8	9	3	2	1			24
	3						3	6	4	1			14
	4							2	2	1	2		7
	5									1	3		4
	6									1		1	2
	Total	4	15	34	53	56	30	19	12	5	5	1	234

* E. Czuber: Die statistischen Forschungsmethoden, Wien, 1921.

TABLE III

x = Diameter of the stem.
y = Length of the longest flower petal in millimeters.
Working mean, $x_m = .825$, $y_m = 34.5$.
Class width of $x = .4$ mm. of $y = 6$ mm.

x	Total times x	P.S. times x	y	Total times y	P.S. times y
−4	16	12	−4	4	4
−3	45	45	−3	21	21
−2	68	68	−2	60	58
−1	53	45	−1	45	33
0	(182)	(170)	0	(130)	(116)
1	30	6	1	48	8
2	38	4	2	48	2
3	36	0	3	42	
4	20	0	4	28	
5	25	0	5	20	
6	6	0	6	12	
	(155)	(10)		(198)	(10)
Mean	−27			+68	

The P.S. columns are the partial sums as explained in the previous table. The work of multiplying the totals by the class marks and of adding them has been separated here from the table.

We obtain $N = 234$, $N_{-x} = 106$, $N_{-y} = 135$

$$r_1 = \frac{170 - 10 - \frac{27}{234} \times 135}{130 + \frac{68}{234} \times 135} = .805$$

$$r_2 = \frac{116 - 10 + \frac{68}{234} \times 106}{182 - \frac{27}{234} \times 106} = .834$$

$$r = .82$$

Pearson's coefficient for this table is $r = .83$.

Finally we illustrate by a small non-grouped table where the partial sums can be written down immediately.

TABLE IV

Correlation between Ages of Husband and Wife

Age of Husband	Age of Wife	Deviation Husband	Deviation Wife
22	18	−8	−8
24	20	−6	−6
26	20	−4	−6
26	24	−4	−2
27	22	−3	−4
27	24	−3	−2
28	27	−2	+1
28	24	−2	−2
29	21	−1	−5
30	25	0	−1
30	29	0	+3
30	32	0	+6
31	27	+1	+1
32	27	+2	+1
33	30	+3	+4
34	27	+4	+1
35	30	+5	+4
35	31	+5	+5
36	30	+6	+4
37	32	+7	+6
Ave 30	26		

Here 0-deviations occur in the third column. Hence[5]

$$\underset{+x}{S}y = 26 + \tfrac{1}{2} \times 8 = 30, \qquad \underset{+x}{S}x = 33, \qquad \underset{+y}{S}x = 31, \qquad \underset{+y}{S}y = 36,$$

$$r_1 = .86, \qquad r_2 = .91, \qquad r = .88 \text{ (Pearson's } r = .86)$$

Appendix

Proof of formula (1), page 1. The following notations will be used:

$(f(x))^0$ = probable value of $f(x)$

$(f(y))^0_x$ = probable value of $f(y)$ for a fixed x.

$$sgx = \text{sign of } x = \frac{x}{|x|} \text{ for } x \neq 0. \quad sgx = \begin{matrix} +1 \\ 0 \\ -1 \end{matrix} \text{ if } x \gtreqless 0.$$

[5] See page 7.

The assumption of linear regression means that

$$y_x^0 - y^0 = r_{y:x}(x - x^0) \tag{4}$$

We multiply both sides of (4) by some arbitrary function $\phi(x)$ of x and get

$$(y_x^0 - y^0)\phi(x) = r_{y:x}(x - x^0)\phi(x).$$

Both sides are functions of x. We shall take their probable values for all x's.

Now, for a fixed x, $y_x^0\phi(x) = (y\phi(x))_x^0$ and the probable value of $(y\phi(x))_x^0$ for all x's is equal to the total probable value $(y\phi(x))^0$. So we have

$$(y\phi(x))^0 - (y^0\phi(x))^0 = r_{y:x}((x - x^0)\phi(x))^0$$

$$r_{y:x} = \frac{((y - y^0)\phi(x))^0}{((x - x^0)\phi(x))^0} \tag{5}$$

If now we take x^0y^0 as the origin, we get

$$r_{y:x} = \frac{(y\phi(x))^0}{(x\phi(x))^0}$$

and similarly

$$r_{x:y} = \frac{(x\phi_1(y))^0}{(y\phi_1(y))^0}$$

where ϕ_1 is another arbitrary function.

Replacing the probable values by the respective arithmetic means we get

$$r_{y:x} = \frac{Sy\phi(x)}{Sx\phi(x)} \quad \text{and} \quad r_{x:y} = \frac{Sx\phi_1(y)}{Sy\phi_1(y)} \tag{6}$$

with $\bar{x}$, $\bar{y}$ as the origin.

By a suitable choice of the still arbitrary functions ϕ and ϕ_1, we may derive all the various expressions for regression coefficients. Taking, for instance, $\phi(x) = x$, $\phi_1(y) = y$, we get Pearson's expressions. Taking $\phi(x) = sg(x - \alpha_1)$, $\phi_1(y) = sg(y - \alpha_2)$, α_1 and α_2 being constants, we have

$$r_{y:x} = \frac{Sy\, sg(x - \alpha_1)}{Sx\, sg(x - \alpha_1)}, \qquad r_{x:y} = \frac{Sx\, sg(y - \alpha_2)}{Sy\, sg(y - \alpha_2)} \tag{7}$$

and if we make $\alpha_1 = \alpha_2 = 0$

$$r_{y:x} \quad \frac{Sy\, sg\, x}{Sx\, sg\, x}, \qquad r_{x:y} = \frac{Sx\, sg\, y}{Sy\, sg\, y} \tag{8}$$

Since $Sx = Sy = 0$, we can add Sy or Sx to the numerators and denominators. Adding Sy to the numerator, Sx to the denominator and multiplying both sides of the fraction by $\frac{1}{2}$ we get

$$r_{y:x} = \frac{\frac{1}{2}Sy(sg(x - \alpha_1) + 1)}{\frac{1}{2}Sx(sg(x - \alpha_1) + 1)} \tag{9}$$

Instead of (9) we can write

$$(10)\qquad r_{y:x} = \frac{\underset{x>\alpha_1}{S}\,y + \tfrac{1}{2}\underset{x=\alpha_1}{S}\,y}{\underset{x>\alpha_1}{S}\,x + \tfrac{1}{2}\underset{x=\alpha_1}{S}\,x}$$

since the operations of (9) multiply the y ordinates by 0, $\frac{1}{2}$, 1 according as the x's are $\lesseqgtr \alpha_1$.

The expression (10), with a suitable choice of α_1 should be used for the purpose of numerical calculation of r. For instance, when calculating r from the data of Table IV, we took $\alpha_1 = \alpha_2 = 0$ and had

$$r_{y:x} = \frac{\underset{+x}{S}y + \tfrac{1}{2}\underset{x=0}{S}\,y}{\underset{+x}{S}x}$$

When dealing with data which are arranged in a grouped table (Tables I and II) we take α_1 equal to the x-ordinate of that classline which is nearest to the mean. $\left(\text{In Table I } \alpha_1 = .5 - \frac{480.5}{1376}\right)$.[6] With that choice of α_1 the sums $\underset{x=\alpha_1}{S}$ disappear and the sums $\underset{x>\alpha_1}{S}$ are equivalent to the corresponding sums $\underset{+x}{S}$. Hence we have

$$(11)\qquad r_{y:x} = \frac{\underset{+x}{S}y}{\underset{+x}{S}x} \quad \text{and similarly} \quad r_{x:y} = \frac{\underset{+y}{S}x}{\underset{+y}{S}y}$$

Instead of (9) we can also write

$$(9a)\qquad r_{y:x} = \frac{\tfrac{1}{2}Sy(sg(x-\alpha_1)-1)}{\tfrac{1}{2}Sx(sg(x-\alpha_1)-1)}$$

This leads to

$$(11a)\qquad r_{y:x} = \frac{\underset{-x}{S}y}{\underset{-x}{S}x} \quad \text{and} \quad r_{x:y} = \frac{\underset{-y}{S}x}{\underset{-y}{S}y}$$

[6] It is desirable to chose the absolute values of the α's small so that the maximum number of data enter into the calculation of r. However, to take $\alpha_1 = \alpha_2 = 0$ would necessitate a division of the middle arrays of a grouped table, a laborious process. Hence the choice of the α's as described above.

Proof of the standard deviations of Formula (2).

In my article on standard deviations and correlations of moments[7] the standard deviations of the expressions used in this article have been derived.

In the following, the notation of the Metron article just referred to will be used. We use the symbols:

$$P_{m,n} = \sum x^m \cdot y^n$$
$$P_{/m,n} = \sum x^m sgxy^n$$
$$P_{m,/n} = \sum x^m y^n sgy$$
$$P_{/m,/n} = \sum x^m sgxy^n sgy$$

The summations indicated extend over all observations. The true or probable values of the same expressions are indicated by using p instead of P.

$$r_{x:y} = r_1 = \frac{P_{1/0}}{P_{0/1}}$$

We derive the standard deviations by defining the deviations as first variations.

$$\log r_1 = \log P_{1/0} - \log P_{0/1}$$

$$\frac{\delta r_1}{r_1'} = \frac{\delta P_{1/0}}{p_{1/0}} - \frac{\delta P_{0/1}}{p_{0/1}}$$

(12) $$\sigma r_1^2 = [(\delta r_1)^2]^0 = (r_1')^2 \left[\left(\frac{\delta P_{1/0}}{p_{1/0}} - \frac{\delta P_{0/1}}{p_{0/1}}\right)^2\right]^0$$

The probable values of the terms on the right hand side of the last equation are derived on pages 17–19 and listed on pages 32–33 of the Metron article referred to. The proofs which imply essentially a process of variation of Stieltje's integrals will not be given here. From pages 32–33 we take

(13)
$$[(\delta P_{1/0})^2]^0 = \frac{p_{20} - p_{1/0}^2}{N}, \qquad [(\delta P_{0/1})^2]^0 = \frac{p_{02} - p_{0/1}^2}{N}$$
$$[(P_{1/0}\delta P_{0/1}]^0 = \frac{p_{11} - p_{1/0}p_{0/1}}{N}$$

so that

(14) $$\sigma_{r_1}^2 = \frac{1}{N}(r_1')^2\left[\frac{p_{20}}{p_{1/0}^2} + \frac{p_{02}}{p_{0/1}^2} - \frac{2p_{11}}{p_{1/0}p_{0/1}}\right]$$

Assuming Gaussian distribution, we can put

$$p_{20} = \frac{\pi}{2}p_{/10}^2 \qquad p_{02} = \frac{\pi}{2}p_{0/1}^2 \qquad p_{11} = r\sqrt{p_{02}p_{20}} = r\frac{\pi}{2}p_{/10}p_{0/1}$$

[7] Felix Bernstein: "Die mittleren Fehlerquadrate und Korrelationen der Potenzmomente und ihre Anwendung auf Funktionen der Potenzmomente," Metron, Vol. X, N. 3 Nov. 1932).

Hence

$$(15) \qquad \sigma_{r_1}^2 = \frac{1}{N}\cdot\frac{\pi}{2}(r_1')^2\left(1 + \frac{p_{/10}^2}{p_{1/0}^2} - 2r\frac{p_{/10}}{p_{1/0}}\right)$$

Replacing the theoretical values by their corresponding empirical values, we have

$$(16) \qquad \sigma_{r_1}^2 = \frac{\pi r_1^2}{2N}(1 + m^2 - 2rm) \qquad \text{where } m = \frac{Sx\, sg\, x}{Sx\, sg\, y}.$$

The formula for $\sigma_{r_1}^2$ has been derived here for the value of r_1 as given by (8) i.e. $r_1 = \dfrac{Sx\, sg\, y}{Sy\, sg\, y}$. In fact, we used $r_1 = \dfrac{Sx\, sg\,(y-\alpha)}{Sy\, sg\,(y-\alpha)}$ in the examples in the article, and α had some value absolutely smaller than .5. To use equation (16) for the standard deviation of r_1 is within the limits of the required degree of accuracy; hence we shall disregard the difference. In a later paper the standard deviation of r_1 for any α will be derived by using the method described in the Metron article, for a different purpose.

To prove the statement in the footnote to page 7

To find the value of r_2 that makes

$$Sf(x)\,(y - r_2 x)^2 \text{ a minimum.}$$

By differentiating we get

$$Sf(x)(y - r_2 x)\,x = 0$$

$$r_2 = \frac{Sxf(x)y}{Sxf(x)x}$$

If $f(x) = 1$ we get Pearson's coefficient.

If $f(x) = \dfrac{1}{|x|}$ $(x \neq 0)$ we get

$$r_2 = \frac{S\dfrac{x}{|x|}y}{S\dfrac{x}{|x|}x} = \frac{Sy\, sg\, x}{Sx\, sg\, x}$$

New York University,
Departments of Anatomy of the Graduate School and the College of Dentistry.

METHODS OF OBTAINING PROBABILITY DISTRIBUTIONS[1]

By Burton H. Camp

The emphasis of this paper will be on method. Special results will be cited in order to illustrate the methods rather than to summarize achievement in the field; for that has been done already by Rider (1930, 1935) Irwin (1935) and Shewhart (1933) in recent surveys. The purpose is to describe and to illustrate most of the methods that have been used to determine exact probability distributions, and to show that they are all derivable from one fundamental theorem. In order to prove this unity in a simple manner, it will be desirable to omit from consideration methods which are essentially ingenious forms of counting, such as are used in sampling without replacements from finite universes, and in finding the sampling distribution of a percentile.

The general problem to be discussed may be stated as follows: N individuals $(t_1, \cdots, t_N)$ are drawn, one at a time with replacements, from a universe whose probability distribution is $\phi(t)$. A certain single valued function of the t's is formed. This is called a parameter of the sample, and is frequently also, but not necessarily, a useful estimate of the corresponding parameter of the universe. The problem is to find its probability distribution, $f(x)$. As usual, a probability distribution is a function which is required to be defined, except perhaps at a set of measure zero, throughout the infinite domain of its variables; it is nowhere negative, and its integral over its domain is unity.

Most of the more recent developments of the theory relate to a more general form of this problem. Instead of N individuals, there are N sets of n individuals in each set, and these sets are drawn respectively from $M(M \leq N)$ universes, each of which is described by a function of n independent variables, thus:

$$\phi^{(i)}(t_1, \cdots, t_n); \; (i = 1, \cdots, M). \tag{1}$$

Instead of a single parameter there are P parameters, and each is a single valued function of the observed values of the nN individuals in the sample, thus:

$$x_i = g_i(t_1^{(1)}, \cdots, t_n^{(1)}; \cdots; t_1^{(N)}, \cdots, t_n^{(N)}); \; (i = 1, \cdots, P) \tag{2}$$

The first method to be described is fundamental and will be designated as

Theorem I. Let it be required that each g as described in (2) be not only single valued but also constant at most in a set of measure zero in the nN-way space of the t's. Then

$$\int_p f(x_1, \cdots, x_P)\, dX = \int_q \phi(t_1^{(1)}, \cdots, t_n^{(N)})\, dT \tag{I}$$

[1] Presented to the American Mathematical Society at a meeting devoted to expository papers on the theory of statistics, April 11, 1936.

where X is the space of x's and T the space of the t's, p is any measurable set of points in X, and q is the set in T for which g is in p. Often p is the P dimensional cube $(x_i + \Delta x,\ i = 1, \cdots P)$ at the point $(x_1, \cdots, x_p)$ and then q is the set where

$$x_i \leqq g_i \leqq x_i + \Delta x; \qquad (i = 1, \cdots, P) \tag{3}$$

and ϕ is the simultaneous distribution of the sets of t's,

$$\phi^{(1)}(t_1^{(1)}, \cdots, t_n^{(1)}) \cdots \phi^{(N)}(t_1^{(N)}, \cdots, t_n^{(N)}). \tag{4}$$

In this $\phi^{(i)}$ is the universe from which the $t^{(i)}$ set of t's is drawn. Obviously, if $N > M$, some of the $\phi^{(i)}$'s are identical, and then it is assumed that the several sets are drawn independently. Often, all of the N sets of t's are drawn from the same universe. Then $M = 1$ and all these ϕ's are identical, and (4) becomes

$$\phi = [\phi^{(1)}(t_1^{(1)}, \cdots, t_n^{(1)}] \cdots [\phi^{(1)}(t_1^{(N)}, \cdots, t_n^{(N)})].$$

In the special case where there is but one parameter ($P = 1$) and but one individual in the sample ($n = N = 1$), and p is an interval, formula (I) becomes

$$\int_x^{x+\Delta x} f(x)\, dx = \int \phi dt; \tag{Ia}$$

and in the very special case where it is also true that q is an interval it becomes

$$f(x) = \phi(t) \quad \frac{dt}{dx} \tag{Ib}$$

provided also that certain derivatives (to be specified later in the proof) exist, where t is now the inverse solution of the equation,

$$x = g(t). \tag{5}$$

The proof of formula (I) is immediate, if one is willing to assume the existence of the probability distribution f; for then the left side is by definition the probability that the x's lie in p, and this is also the meaning of the right side of (I). (Ia) can be proved without assuming initially the existence of $f(x)$, for then the existence of $f(x)$ can be inferred from the existence of the right side of (Ia), because $f(x)$ may be set equal (except perhaps at a set of measure zero) to the upper right hand derivative, with respect to Δx (Δx is a variable, and x is fixed), of $\int_q \phi\, dt$, provided that one adds the condition that this derivative is nowhere infinite. The point at issue here is merely the existence of a primative for a monotone increasing function of Δx. (Ib) may be derived from (Ia) by taking the derivative of both sides with respect to Δx, if the derivatives are continuous.

Theorem I, in these various forms is used a great deal, especially in the last form (Ib). This affords one freedom to choose the most desirable function for purposes of tabulation. R. A. Fischer's z distribution, a logarithm, is an important illustration. Many authors have been interested in so choosing the

function that its distribution shall be normal. They include several of the older writers, and more recently H. L. Rietz (1921, 1927), and G. A. Baker (1932, 1934). However, the theorem is of special importance in the theory, for all the other principal methods of obtaining probability distributions are essentially corollaries of it. These corollaries will be called Theorems II, III, and IV.

THEOREM II. Let $\bar{p}$ (the measure of p) and $\bar{q}$ (the measure of q) be infinitesimals of the same order and let both the oscillation of f(*i.e.* maximum f- minimum f) in p and the oscillation of ϕ in q be infinitesimals; then (I) may be written,

$$f\bar{p} = \phi\bar{q}, \tag{II}$$

where f applies to any point of p and ϕ to the corresponding point of q. This equation (II) is an approximate equation in the sense that differences of higher order than those retained are neglected. In particular, with the conditions used in formula (Ia), equation II becomes

$$f\Delta x = \phi\bar{q}.$$

The left side of (II) is an approximation to the probability sought. The right side shows that, in order to evaluate it, one need only find the volume in T space of the differential element q and multiply it by the value of ϕ in q. Formula (II) expresses the so-called geometrical method used by many authors, *e.g.*, by R. A. Fisher (1915, 1925), by Wishart (1928), and by Hotelling (1925, 1927). The chief difficulty in connection with it is in finding the volume of nN-dimensional q. In order to display the advantages and disadvantages of this method we shall pause at this point and look at a concrete example.[2]

Let two individual (t_1, t_2) be drawn independently from a normal universe and consider the simultaneous distribution $f(x, y)$ of the sum, $x = t_1 + t_2$, and product, $y = t_1t_2$, the mean of the universe being chosen as the origin. Here $N = 2$, $n = 1$, $M = 1$, and so,

$$\phi = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(t_1^2 + t_2^2)} = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x^2 - 2xy)}. \tag{6}$$

The point set q is the area lying between the two adjacent hyperbolae,

$$t_1t_2 = y, \qquad t_1t_2 = y + \Delta y,$$

and also between the two adjacent lines,

$$t_1 + t_2 = x, \qquad t_1 + t_2 = x + \Delta x,$$

where Δx and Δy are infinitesimals and are equal. This area may be computed by simple integration and is:

[2] See also C. C. Craig (1936). Craig uses another method to be explained later (formula IIIa).

$$q = \frac{2\Delta x\,\Delta y}{\sqrt{x^2 - 4y}} \qquad \text{if} \quad x^2 > 4y,$$

$$= 0 \qquad \text{if} \quad x^2 < 4y.$$

Hence II gives us immediately the desired result:

$$f(x, y)\,\Delta x\Delta y = \frac{1}{\pi\sigma^2}\, e^{-\frac{x^2 - 2y}{2\sigma^2}} \frac{1}{\sqrt{x^2 - 4y}} \cdot \Delta x\Delta y, \qquad \text{if} \quad x^2 > 4y,$$

$$= 0 \quad \text{if} \quad x^2 < 4y.$$

If $x^2 = 4y$, $\bar{q}$ is an infinitesimal of lower order than $\bar{p} = (\Delta x)^2$, and so Theorem II does not apply. In this case we must go back to Theorem I, and from that we can learn that the probability,

$$\int_p f\,dx\,dy,$$

is an infinitesimal of the first order if $p = \Delta x\,\Delta y = (\Delta x)^2$ is of the second order. Hence it cannot be approximately represented by a finite number times $\bar{p}$. The oscillation of f in p is infinite. The form of the surface $f(x, y)$ is interesting. The ordinates rise to infinity on the contour of the parabola $x^2 = 4y$, and vanish within it. The surface is symmetrical with respect to the plane $x = 0$, but not with respect to the plane $y = 0$. However, it is clear that the total probability of any given product, y (*i.e.* the probability of this y for all possible values of x), is the same as the total probability of $-y$; hence

$$\int_{-\infty}^{\infty} f(x, y)\,dx = \int_{-\infty}^{\infty} f(x, -y)\,dx,$$

and the corresponding formulae,

$$\frac{2}{\pi\sigma^2}\, e^{\frac{y}{\sigma^2}} \int_{\sqrt{4y}}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \frac{1}{\sqrt{x^2 - 4y}}\,dx \qquad (y > 0),$$

and

$$\frac{2}{\pi\sigma^2}\, e^{\frac{y}{\sigma^2}} \int_0^{\infty} e^{-\frac{x^2}{2\sigma^2}} \frac{1}{\sqrt{x^2 - 4y}}\,dx \qquad (y < 0),$$

must be equal; both may be reduced to the single form

$$F(y) = \frac{1}{\pi\sigma^2} \int_0^{\infty} e^{-\frac{1}{2\sigma^2}\left(t^2 + \frac{y^2}{t^2}\right)} \frac{dt}{t}, \qquad \text{if} \quad y \neq 0.$$

This is the probability distribution of y.

With this example before us, let us now reconsider the theory:

(*i*) The requirement (in II) that the oscillation of ϕ be infinitesimal in q

will be satisfied if one can show that ϕ may be expressed as a continuous function of the parameters $(x_1, x_2, \cdots, x_p)$. In our example these parameters were x and y and ϕ was so expressible (6). But if we had tried initially to find by means of (II) the distribution of the product y, independently of what values x might have, we should have been stopped at this point, because ϕ is not expressible in terms of y alone. We should also have been stopped by the requirement that $\bar{q}$ be infinitesimal of order Δy, for q would have been the space between two hyperbolas and its area for any fixed $(\Delta y > 0)$ would have been infinite. But, when thus stopped at that first point, it would have been clearly indicated to us that the distribution of y might have been found *via* the detour of finding the simultaneous distribution of both x and y, because an attempt to express ϕ in terms of y would have led to the given expression in terms of both x and y. For a similar reason R. A. Fisher (1925) was able to find the distribution of the variance by finding first the simultaneous distribution of the variance and the mean. Also, he was thus able to find the distribution of the coefficient of correlation by finding first the simultaneous distribution of all the first and second order moments.

(*ii*) A distinct advantage of this method is that q is independent of the universe ϕ, so that once found it may be used in connection with any universe which satisfies the condition that it can be expressed as a continuous function of the parameters. Thus, the distribution of the sum and product in our example may equally well be found for the universe described by the Type III curve, $Ate^{-at}(t > 0)$. For, then

$$\phi = A^2 t_1 t_2 e^{-a(t_1+t_2)} = A^2 y e^{-ax},$$

and so, using one-half of the same $\bar{q}$ as before, since now $x, y \geqq 0$,

$$f(x, y) = A^2 y e^{-ax} \frac{2}{\sqrt{x^2 - 4y}} \qquad \text{if} \quad x^2 > 4y,$$

$$= 0 \qquad \text{if} \quad x^2 < 4y.$$

From this, $F(y)$ can be found by integration (*c.f.* Kullbach, 1934)

$$F(y) = A^2 y \int_{\sqrt{4y}}^{\infty} \frac{e^{-ax}}{\sqrt{x^2 - 4y}}\, dx = \frac{A^2 y}{2} \int_0^{\infty} \frac{e^{-a\left(u + \frac{y}{u}\right)}}{u}\, du.$$

As another illustration, consider a normal universe of n intercorrelated variables in which all the total intercorrelations are equal to r (*e.g.*, the statures of n brothers) and let the sample be a single group of n (one individual for each variable).

$$\phi = \frac{1}{(2\pi)^{n/2} R} e^{-\frac{1}{2R}\left[k_1 \Sigma t_i^2 + k_2 \sum_{i \neq j} t_i t_j\right]},$$

where $R = (1 - r)^{n-1}[1 - (n - 1)r]$, $k_1 = (1 - r)^{n-2}[1 - (n - 2)r]$, and $k_2 = -r(1 - r)^{n-2}$. Suppose one wishes to find the simultaneous distribution

of the variance x and the mean y for such samples.[3] Since for Student's problem Fisher has found the value of q for this x and y to be

$$\bar{q} = cx^{\frac{n-3}{2}} \Delta x \Delta y,$$

their distribution $f(x, y)$ for this universe may be written down immediately. In terms of x and y the bracket in the exponent of ϕ is $y^2(k_1 n - k_2 n + k_2 n^2) + xn(k_1 - k_2)$, and so $f(x, y)$ is the product of $\bar{q}$ and this form of ϕ:

$$f(x, y) = K e^E x^{\frac{n-3}{2}}, \qquad E = -\frac{1}{2R}[(k_1 n - k_2 n + k_2 n^2)y^2 - n(k_1 - k_2)x].$$

(*iii*) Another attribute of this method is that it sometimes lends itself to easy extensions from a simple case where there is only one restriction ($N - 1$ degrees of freedom) to similar cases when there are more restrictions. Thus R. A. Fisher (1924) proceeded from the variance of a sample from a single universe to the variance from a set of universes, as required in the theory of analysis of variance; and thus also (1915) he had proceeded from the distribution of r to that of multiple R; and Hotelling (1927) showed how these distributions could be obtained when the values of each variate were themselves intercorrelated (as in a time series) and not merely correlated with values of the other variates.

THEOREM III. Now let us consider again the fundamental form (I). For convenience let $nN = m$. If the conditions will not permit us to write the right side in the form in (II), it is still possible that we may be able to find that $(m + 1)$-dimensional volume by some other method. In particular, whenever it is possible to iterate the integral once we have the formula:

$$\text{(III)} \qquad \int_p f dX = \int_{T'} dT' \int_{q_m} \phi \, dt_m,$$

where q_m is the section of q by t_m space at the point $(t_1, \cdots, t_{m-1})$ of T' space, T' space being the space of the $(t_1, \cdots, t_{m-1})$ coordinates. With added conditions one may deduce from (III), for the case where there is but a single parameter x, the approximate equation:

$$\text{(IIIa)} \qquad f\,dx = dx \int_{T'} dT' \cdot \phi(t_1, \cdots, t_m) \frac{dt_m}{dx},$$

in which t_m is supposed to have been expressed in terms of the other coordinates by solving the equation $x = g(t_1, \cdots, t_m)$. It is an approximate equation in the same sense as (II) was. Sufficient conditions for this change in the left side of (III) have already been mentioned in discussing (II). The propriety of making the corresponding change in the right hand side may be left for determination when the form of ϕ is given. It will perhaps be sufficient here to point out that our earlier example illustrates both the case where this change

[3] A special case of a more general problem solved first by R. A. Fisher.

is permissible and where it is not. For, let it be required to find the distribution $f(y)$ of the product $y = t_1 t_2$ without reference to the sum, $t_1 + t_2$. Formula (III) yields

$$(7) \qquad \int_y^{y+\Delta y} f(y)\,dy = 2\int_0^\infty dt_1 \int_{y/t_1}^{(y+\Delta y)/t_1} dt_2 \cdot \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left(t_1^2 + t_2^2\right)}.$$

This is valid for every value of y including $y = 0$. If $y \neq 0$, we may change the right hand side as in (IIIa) and obtain as the probability that y is in the interval $(y, y + \Delta y)$:

$$(8) \qquad \int_y^{y+\Delta y} f(y)\,dy = \frac{\Delta y}{\pi\sigma^2} \int_0^\infty \frac{1}{t_1} e^{-\frac{1}{2\sigma^2}\left(t_1^2 + \frac{y^2}{t_1^2}\right)} dt_1 + \epsilon,$$

where ϵ is a differential of higher order than Δy. This may be proved by computing the difference between the value of (7) when t_2 has constantly the value $(y + \Delta y)/t_1$ and when it has constantly the value y/t_1. If $y = 0$ this change in the right side of (7) is not valid; it is easily seen that in this case the integral on the right of (8) is infinite. It may be shown, however, in this case that

$$(9) \qquad \int_0^{\Delta y} f(y)\,dy = \tfrac{1}{4} - \frac{1}{2\pi}\int_1^\infty \frac{e^{-\frac{x\Delta y}{\sigma^2}}\,dx}{x\sqrt{x^2 - 1}},$$

and that this is an infinitesimal, and that it is of order as small as one.

Many authors think of (IIIa) as the fundamental formula in the theory of probability distributions. One of the simplest and earliest applications of it was to establish the so-called reproductive property of the normal law: that the sum of two variates is distributed normally if each is distributed normally. Jackson (1935) has used it to establish a similar property for two Type III distributions which have the same exponent of e. Usually this integral is difficult to evaluate when $N > 2$ because of the unsymmetrical form into which it is cast, but when $N = 2$ and there is but one parameter (IIIa) it is perhaps the most convenient of all the formulae.

Theorem IV. An exceedingly useful formula is obtainable from (I) in the following manner. Let $\theta(x_1, \cdots, x_P; \alpha_1, \cdots, \alpha_Q)$ be a finite single valued function of the old parameters (x) and of some new parameters (α). Subject to general conditions to be stated we may write:

$$(\text{IV}) \qquad \int_X \theta f\,dX = \int_T \theta'\phi\,dT,$$

an identity with respect to each α, where θ' is the result of substituting (2) for the x's in θ.

Since this theorem has not been proved in this general form, an outline of the proof will be given. Sufficient conditions are:

(a) All the integrals involved shall exist.

(b) If p is limited (in the sense that it lies within a finite hypersphere), so is q, and conversely.

Proof. Let X_0 be a limited p set and T_0 the corresponding q set such that both (c) and (d) hold ($\epsilon > 0$):

$$\int_{X_0} f\theta\, dX - \int_X f\theta\, dX < \epsilon, \tag{c}$$

$$\int_{T_0} \phi\theta'\, dT - \int_T \phi\theta'\, dT < \epsilon. \tag{d}$$

It is easy to see that such an X_0 and a corresponding T_0 do exist, as follows: Let X_0' be a limited set for which (c) is true, and for which it will remain true no matter what points are added to X_0'. Similarly, let T_0' be a limited set for which (d) is true and for which it will remain true, no matter what points are added to T_0'. Presumably X_0' and T_0' do not correspond to each other, but we may now let X_0 be the totality of all the points of X_0' and of all those points of X corresponding to T_0', and let T_0 be the totality of all the points of T_0' and of all those points of T corresponding to X_0'. Then X_0 and T_0 do correspond to each other and have the desired properties (c) and (d). Now, since θ is finite, it is limited in X_0. Let

$$|\theta| < H \text{ in } X_0. \tag{e}$$

Divide the interval $(-H, H)$ into s equal subintervals of length h, thus defining in X_0 according to Lebesgue the measurable sets,

p_i $(i = 1, \cdots, s)$, and corresponding q_i sets in T_0:

$$\begin{cases} Os\ \theta \leqq h \text{ in } p_i, \\ Os\ \theta' \leqq h \text{ in } q_i. \end{cases} \tag{f}$$

Choose arbitrarily any point of p_i and let k_i be the corresponding value of θ. Then let

$$\bar{\theta} = k_i \text{ in } p_i \ (i = 1, \cdots, s), \text{ and similarly let}$$
$$\bar{\theta}' = k_i \text{ in } q_i \ (i = 1, \cdots, s).$$

Then

$$\int_{X_0} \bar{\theta} f\, dX = \sum_i k_i \int_{p_i} f\, dX,$$

and

$$\int_{T_0} \bar{\theta}'\phi\, dT = \sum_i k_i \int_{q_i} \phi\, dt.$$

Since by (I)

$$\int_{p_i} f\, dX = \int_{q_i} \phi\, dT,$$

$$\int_{X_0} \bar{\theta} f\, dX = \int_{T_0} \bar{\theta}'\phi\, dT. \tag{g}$$

Now

$$(\bar{\theta} - \theta) f \, dX \qquad \int_{X_0} | \bar{\theta} - \theta | f \, dX \leqq h \int_{X_0} f \, dX,$$

and

$$\left. \int_{T_0} (\bar{\theta}' - \theta') \, dX \right| \leqq h \int_{T_0} \phi \, dX.$$

So, as h approaches zero both sides of (g) approach limits and their limits are equal:

$$\int_{X_0} \theta f \, dX = \int_{T_0} \theta' \phi \, dT.$$

Hence by (c) and (d) the integrals

$$\int_X \theta f \, dx, \qquad \int_T \theta' \phi \, dT,$$

differ at most by 2ϵ, and so, being independent of ϵ they do not differ at all.

In order to determine the form of f from (IV) one must first evaluate the right side,

$$\int_T \theta\phi \, dt = \psi(\alpha_1, \cdots, \alpha_q);$$

and then solve the integral equation,

$$\int_X \theta f \, dX = \psi. \tag{10}$$

It is the solution of this equation that usually presents the most difficulty. Particular forms of θ that are being used are

$$\theta = e^{\alpha_1 x + \cdots + \alpha_p x_p}, \tag{11}$$

in which case ψ is said to be the "characteristic function" or "moment generating function"; and

$$\theta = x_1^{\alpha_1} \cdots x_p^{\alpha_p}, \tag{12}$$

in which case ψ is a "moment function" or "moment" of f. Other forms might be used. For example, a very convenient method of demonstrating the correctness of the usual formula for the simultaneous distribution of the correlation (x), means (y, z), and variances (u, v), in samples from a normal bivariate universe is by the use of

$$\theta = e^{\alpha_1 (u^2 + v^2 + y^2 + z^2) + \alpha_2 (uvx + yz)}.$$

This method of finding f is not a final determination of the probability function desired until it has been shown that the solution is unique, a serious problem

in itself; it is one of those which Professor Shohat may consider.[4] There are three methods of solving the integral equation (10):

(*i*) The first might be called guessing. Though unscientific, it is in fact often effective. Especially is it available if the distribution has already been surmised but not demonstrated. Thus, it was open to Student (1908) when he correctly surmised the distribution of the variance. Similarly it was open to Soper (1913) when he incorrectly surmised the distribution of r.

(*ii*) Papers by Romanovsky (1925) and Wilks (1932) have shown how the problem of solving the integral equation may be shifted to the problem of solving a partial differential equation, but this in turn may involve the solution of another equally difficult integral equation in the process or determining the arbitrary function.

(*iii*) If each α be replaced by an imaginary βi and one uses a Fourier transform, one arrives at a set of formulae which are most important. For the case where there is but one x and one β, they may be written:

$$\int_{-\infty}^{\infty} e^{i\beta x} f(x)\, dx = \int_T e^{i\beta g}\, \phi\, dT = \psi(\beta). \tag{13}$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\beta x}\, \psi(\beta)\, d\beta. \tag{14}$$

Dodd (1925) has given an equivalent set of formulae involving only real variables. It is easy to prove that both sets may be changed to the single formula,

$$f(x) = \frac{1}{\pi} \int_T \phi\, dt \int_0 \cos \beta(x - g)\, d\beta. \tag{15}$$

Kullbach (1936) has established the validity of the formulae corresponding to (13) and (14) for the general case of $(P + Q)$ parameters. Wishart and Bartlett (1933) used the general forms to find the distribution of the generalized product moment in samples from an n-dimensional normal system.

When the solution of the integral equations of (IV) cannot be found, one has to put up with the semi-invariants or with the moments of f. Formulae (IV) and (11) yield the semi-invariants, (IV) and (12) the moments about the given origin, and from either of these one may obtain the moments about the mean point. These methods are old but they are still important. Time does not permit me to discuss them, because it would not be proper to close this paper without some reference to limit methods.

Limit Methods. It is well known that the distribution of means of samples taken from almost[5] any universe approaches the normal law as a limit as N becomes infinite. This theorem is subject to great generalizations, as is indicated in papers of A. Liapounoff (1901), S. Bernstein (1926), Romanovsky

[4] In a later paper at the same symposium.

[5] There are exceptions. *E. g.*, means of samples taken from the universe $a/\pi(a + t^2)$ have a distribution identical with the universe itself.

(1929, 1930) and C. C. Craig (1932). Subject to very general conditions it has been shown that: If the characteristic function of one probability distribution contains a parameter and approaches as a limit, uniformly in every finite domain of its variables, the characteristic function of another probability distribution; then the first distribution approaches as a limit the second distribution. Hence S. Bernstein and Romanovsky have shown that: If the universe is an n-way correlation solid of a certain very general type, then the n means obtained by a selection of a sample of N sets of variates, $x_i = \frac{1}{N}(t_{i_1} + \cdots + t_{iN})$, $(i = 1, \cdots, n)$, have a distribution which approaches as a limit a normal correlation solid as N becomes infinite. A similar theorem has been established also in the interesting case of Romanovsky's "belonging coefficients", which include K. Pearson's coefficient of racial likeness. Also, by the method of maximum likelihood, Hotelling (1930) has proved that under certain general conditions all optimum estimates of the parameters of a frequency distribution have a joint distribution approaching the normal as N becomes infinite. The validity of the method of maximum likelihood when used for this purpose has been established by J. L. Doob (1934).

Finally, one may note an apparently new limit theorem of another type. Its general nature will be obvious from the following application:

Let a sample of N be drawn from the universe,

$$\begin{aligned} \phi &= Ae^{-at^{2\lambda}}, \quad \text{if} \quad t > 0, \\ &= 0 \qquad\qquad \text{if} \quad t \leqq 0. \end{aligned}$$

It is readily proved, by means of (IV), that the distribution $f(x)$ of the parameter,

$$x = (t_1^{2\lambda} + \cdots + t_N^{2\lambda})^{1/N}$$

is a curve of the form,

$$\begin{aligned} f(x) &= Bx^{N-1}\, e^{-x^{2\lambda}} \text{ where } x > 0, \\ &= 0 \qquad\qquad \text{elsewhere.} \end{aligned}$$

Now let λ become infinite. The universe approaches as a limit the rectangle:

$$\begin{aligned} \Phi &= A \text{ where } 0 \leqq t < 1, \\ &= 0 \qquad\qquad \text{elsewhere.} \end{aligned}$$

The parameter x approaches as a limit X, where X = maximum t_i. The distribution $f(x)$ approaches as a limit the new distribution,

$$\begin{aligned} F(X) &= NX^{N-1} \text{ where } 0 < |X| < 1, \\ &= 0 \qquad\qquad \text{elsewhere.} \end{aligned}$$

Hence we have proved in a new way, what was already known: that the distribution of the greatest variate obtained by sampling from a rectangular universe is of the form $F(X)$.

The limit theorem implicit in this illustration can be established in sufficient generality, but I do not yet know whether it has other applications of value.

REFERENCES

G. A. BAKER, Transformations of bimodal distributions, Annals of Mathematical Statistics, vol. 1 (1932), pp. 334–344.

G. A. BAKER, Transformation of non-normal frequency distributions into normal distributions, Annals of Mathematical Statistics, vol. 5 (1934), pp. 113–123.

S. BERNSTEIN, Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dependantes, Mathematische Annalen, vol. 97 (1926), pp. 1–59.

C. C. CRAIG, On the composition of dependent elementary errors, Annals of Mathematics, vol. 33 (1932), pp. 184–206.

C. C. CRAIG, On the frequency function of xy, Annals of Mathematical Statistics, vol. 7 (1936), pp. 1–15.

E. L. DODD, The frequency law of a function of variables with given frequency laws, Annals of Mathematics, vol. 27 (1925), pp. 12–20.

J. L. DOOB, Probability and statistics, Transactions of the American Mathematical Society, vol. 36 (1934), pp. 759–775.

R. A. FISHER, Frequency distributions of the values of the correlation coefficient in samples from an indefinitely large population, Biometrika, vol. 10 (1915), pp. 507–521.

R. A. FISHER, On a distribution yielding the error function of several well known statistics, Proceedings of the International Mathematical Congress, Toronto, 1924, vol. 2, pp. 805–813.

R. A. FISHER, Applications of "Student's" distribution, Metron, vol. 5, no. 3, (1925), pp. 90–104.

H. HOTELLING, The distribution of correlation ratios calculated from random data, Proceedings of the National Academy of Sciences, vol. 11, no. 10 (1925), pp. 657–662.

H. HOTELLING, An application of analysis situs to statistics. Bulletin of the American Mathematical Society, vol. 33 (1927), pp. 467–476.

H. HOTELLING, The consistency and ultimate distribution of optimum statistics, Transactions of the American Mathematical Society, vol. 32 (1930), pp. 847–859.

J. O. IRWIN, Recent advances in mathematical statistics, Journal of the Royal Statistical Society, vol. 98, part 1 (1935), pp. 88–92.

D. JACKSON, Mathematical principles in the theory of small samples, American Mathematical Monthly, vol. 42 (1935), pp. 344–364.

S. KULLBACH, An application of characteristic functions to the distribution problem of statistics, Annals of Mathematical Statistics, vol. 5 (1934), pp. 263–307.

S. KULLBACH, On certain distribution theorems of statistics, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 407–410.

A. LIAPOUNOFF, Nouvelle forme du théorème sur la limite de probabilité, Memoires de l'Académie de St. Pétersbourg (8), vol. 11 (1901).

P. R. RIDER, A survey of the theory of small samples, Annals of Mathematics, vol. 31 (1930), pp. 577–628.

P. R. RIDER, Recent progress in statistical method, Journal of the American Statistical Association, vol. 30 (1935), pp. 58–88.

H. L. RIETZ, Frequency distributions obtained by certain transformations of normally distributed variates, Annals of Mathematics (2) vol. 23 (1921–22), pp. 292–300.

H. L. RIETZ, On certain properties of frequency distributions of the powers and roots of the variates of a given distribution, Proceedings of the National Academy of Sciences, vol. 13 (1927), pp. 817–820.

U. ROMANOVSKY, On the moments of the standard deviations and of correlation coefficients in samples from a normal population, Metron, Vol. 5(4) (1925), pp. 1–45.

U. ROMANOVSKY, Sur une extension du théorème de A. Liapounoff sur la limite de probabilité, Bulletin de l'Academie des Sciences de l'U. S. S. R. (1929), pp. 209–225.

U. ROMANOVSKY, On the moments of means of functions of one or more random variables, Metron, vol. 8(1-2) (1930), pp. 251-291.

W. A. SHEWHART, Annual Survey of Statistical Technique: Sample Theory, Econometrica, vol. 1 (1933), pp. 225–237.

H. E. SOPER, On the probable error of the correlation coefficient to a second approximation, Biometrika, Vol. 9 (1913), pp. 91–115.

STUDENT, The probable error of the mean, Biometrika, vol. 6 (1908), pp. 1–25.

S. S. WILKS, On the distributions of statistics in samples from a normal population of two variables with matched sampling of one variable, Metron, vol. 9(3–4) (1932), pp. 87–126.

J. WISHART, The generalized product moment distribution in samples from a normal multivariate population, Biometrika, vol. 20A, (1928), pp. 32–52.

J. WISHART AND M. S. BARTLETT, The generalized product moment distribution in a normal system, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 260–270.

WESLEYAN UNIVERSITY.

MOMENT RECURRENCE RELATIONS FOR BINOMIAL, POISSON AND HYPERGEOMETRIC FREQUENCY DISTRIBUTIONS[1]

By John Riordan

1. **Introduction.** This paper gives the development of recurrence relations for moments about the origin and mean of binomial, Poisson, and hypergeometric frequency distributions from the basis of the moment arrays defined by H. E. Soper.[2] This procedure has the advantage of expressing the moments in terms of coefficients which are alike for the three distributions and are derivable by a single process, thus providing a degree of formal coordination of the distributions. For both kinds of moments, the coefficients satisfy relatively simple recurrence relations, the use of which leads to recurrence relations for the moments, thus unifying the derivation of these relations for the three distributions. The relations derived in this way for the hypergeometric distribution are apparently new. Apparently new recurrence relations for certain auxiliary coefficients in the expression of the moments about the mean of binomial and Poisson distributions are also given.

This course of development involves repetition of a number of well-known results which is justified, it is hoped, by the unification obtained.[3]

[1] Presented to the American Mathematical Society, Sept. 3, 1936.

[2] *Frequency Arrays*, Cambridge, 1922.

[3] The following bibliography is taken from a paper *On the Bernoulli Distribution*, Solomon Kullback, Bull. Am. Math. Soc., **41**, 12, pp. 857–864, (Dec., 1935):

A. Fisher, *The Mathematical Theory of Probabilities*, 2d ed., p. 104 ff.

H. L. Rietz, *Mathematical Statistics*, 1927, p. 26 ff.

V. Mises, *Wahrscheinlichkeitsrechnung*, 1931, pp. 131–133.

Risser and Traynard, *Les Principes de la Statistique Mathématique*, 1933, pp. 39–40 and 320–321.

V. Romanovsky, *Note on the moments of the binomial* $(q + p)^n$ *about its mean*, Biometrika, vol. 15 (1923), pp. 410–412.

A. T. Craig, *Note on the moments of a Bernoulli distribution*, Bull. Am. Math. Soc., vol. 40 (1934), pp. 262–264.

A. R. Crathorne, *Moments de la binomiale par rapport à L'origine*, Comptes Rendus, vol. 198 (1934), p. 1202;

A. A. K. Aygangar, *Note on the recurrence formulae for the moments of the point binomial*, Biometrika, vol. 26 (1934), pp. 262–264.

To this, besides Soper's tract already mentioned, should be added:

Ch. Jordan, *Statistique Mathematique*, Paris, 1927.

K. Pearson, *On Certain Properties of the Hypergeometric Series . . .*, Phil. Mag., **47**, pp. 236–246 (1899).

2. Moment Arrays. As developed by Soper, frequency distributions may be exhibited by frequency arrays, in the case of a single variate, in the form:

$$f(A) = \sum_x p_x A^x \tag{2.1}$$

where p_x are the frequencies with which the measures, x, of the character, A, occur in a population.

The substitution $A = e^\alpha$ leads to the moment about the origin array:

$$\begin{aligned} f(e^\alpha) &= \sum_x p_x e^{x\alpha} \\ &= \sum_x p_x \left(1 + x\alpha + \frac{x^2\alpha^2}{2!} + \cdots\right) \\ &= \sum_s m_s \frac{\alpha^s}{s!} \end{aligned} \tag{2.2}$$

where

$$m_s = \sum_x p_x x^s$$

The symbol α is a logical or umbral symbol serving merely to identify the moments in the expansion of the array.

The moment array for moments about the mean is found from the relation:

$$\begin{aligned} \phi(e^\alpha) &= e^{-m_1\alpha} f(e^\alpha) \\ &= \sum_s \mu_s \alpha^s/s! \end{aligned}$$

where m_1 is the first moment about the origin.

The moment arrays for the distributions concerned are as follows:

Binomial $\quad f(e^\alpha) = [1 + p(e^\alpha - 1)]^n = \sum_{x=0}^{n} \binom{n}{x} p^x(e^\alpha - 1)^x$

Poisson $\quad f(e^\alpha) = e^{a(e^\alpha - 1)} = \sum_{x=0}^{\infty} \frac{a^x(e^\alpha - 1)^x}{x!}$

Hypergeometric $\quad f(e^\alpha) = \sum_{x=0}^{\infty} \frac{(l)_x(r)_x}{(n)_x} \frac{(e^\alpha - 1)^x}{x!}$

where the parameters p, n, and a for the binomial and Poisson have the usual significance. The parameters for the hypergeometric distribution, with the substitution $r = s$, follow Soper; Pearson (loc. cit.) uses q, r, n, where $q = l/n$. The notation $(l)_x$ means

$$(l)_x = l(l - 1) \cdots (l - x + 1).$$

It will be seen that, with the usual interpretation of $\binom{n}{x}$ as zero for $x > n$,

the three distributions so far as concerns α may be exhibited by a function of the form

$$f(e^{\alpha}) = \sum_{x=0}^{\infty} A_x(e^{\alpha} - 1)^x$$

where A_x of course depends on the distribution concerned.

3. **Moments About the Origin.** The moments about the origin can then be defined by the equation:

$$\sum_{s=0}^{\infty} m_s \frac{\alpha^s}{s!} = \sum_{x=0}^{\infty} A_x(e^{\alpha} - 1)^x \tag{3.1}$$

and

$$\sum_{x=0}^{\infty} A^x(e^{\alpha} - 1)^x = \sum_{x=0}^{\infty} A_x \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} e^{v\alpha}$$
$$= \sum_{s=0}^{\infty} \frac{\alpha^s}{s!} \sum_{x=0}^{s} x!\, A_x S_{x,s}$$

where $S_{x,s}$ is a Stirling number of the second kind, as used by Jordan (loc. cit.) and defined by

$$x!\, S_{x,s} = \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} v^s = \Delta^x 0^s,$$

$\Delta^x 0^s$ being in the language of the finite difference calculus, a "difference of nothing" that is $\Delta^x n^s \mid n = 0$.

The internal series terminates at s because $S_{x,s} = 0$, $x > s$, as is readily apparent in the finite difference expression. Further $S_{0,s} = 0$, $s \neq 0$; $S_{0,0} = 1$.

By equating coefficients in equation (3.1), m_s, the sth moment about the origin, is given by

$$m_s = \sum_{x=0}^{s} x!\, A_x S_{x,s}. \tag{3.2}$$

The particular forms for the three distributions are as follows:

$$m_s = \sum_{x=0}^{s} (n)_x p^x S_{x,s} \qquad \textit{Binomial} \tag{3.3}$$

$$m_s = \sum_{x=0}^{s} a^x S_{x,s} \qquad \textit{Poisson} \tag{3.4}$$

$$m_s = \sum_{x=0}^{s} \frac{(l)_x (r)_x}{(n)_x} S_{x,s} \qquad \textit{Hypergeometric} \tag{3.5}$$

The Stirling numbers have the following recurrence relation (Jordan loc. cit.):

$$S_{x,s+1} = xS_{x,s} + S_{x-1,s}. \tag{3.6}$$

This relation in conjunction with equations (3.3)–(3.5) leads to moment recurrence relations. The procedure is illustrated for the binomial distribution as follows:

$$\begin{aligned} m_{s+1} &= \sum_{x=0}^{s+1} (n)_x\, p^x\, S_{x,\,s+1} \\ &= \sum_{x=0}^{s+1} (n)_x\, p^x\, (x\, S_{x,\,s} + S_{x-1,\,s}) \\ &= p\, D_p\, m_s + (npm_s - p^2\, D_p\, m_s) \\ &= npm_s + pq\, D_p\, m_s \end{aligned}$$

where $q = 1 - p$.

The steps in the process are expanded as follows:

$$\begin{aligned} \sum_{x=0} (n)_x p^x\, x\, S_{x,\,s} &= \sum_{x=0} (n)_x\, p^x\, x\, S_{x,\,s} \\ &= \sum_{x=0}^{s} (n)_x\, S_{x,\,s}\, pD_p(p^x) \\ &= pD_p\, m_s \end{aligned}$$

$$\begin{aligned} \sum_{x=0}^{s+1} (n)_x\, p^x\, S_{x-1,\,s} &= \sum_{x=0}^{s+1} (n - x + 1)\, (n)_{x-1}\, p^x\, S_{x-1,\,s} \\ &= n \sum_{x=1}^{s} (n)_x\, p^{x+1}\, S_{x,\,s} - \sum_{x=1}^{s} x(n)_x\, p^{x+1}\, S_{x,\,s} \\ &= np\, m_s - p^2 D_p\, m_s \end{aligned}$$

The results for the three distributions are as follows:

(3.7) $$m_{s+1} = npm_s + pqD_pm_s \qquad \textit{Binomial}$$

(3.8) $$m_{s+1} = am_s + aD_am_s \qquad \textit{Poisson}$$

(3.9) $$m_{s+1} = \frac{lr}{n}\, m_s(l - 1, r - 1, n - 1) - (n + 1)\Delta_n m_s \qquad \textit{Hypergeometric}$$

Here D_p and D_a denote differentiation with respect to p and a, respectively, and Δ_n denotes the difference operation with respect to n. For the hypergeometric distribution the moments are functions of l, r, and n as well as of s; $m_s(l - 1, r - 1, n - 1)$ is the same function of $l - 1$, $r - 1$ and $n - 1$ as $m_s(l, r, n)$ is of l, r, n. Equation (3.9) appears to be new.

For convenience of reference, a short table of the Stirling numbers of the second kind follows:

$s \backslash x$	0	1	2	3	4	5
0	1					
1	0	1				
2	0	1	1			
3	0	1	3	1		
4	0	1	7	6	1	
5	0	1	15	25	10	1

$S_{x,s}$

4. Moments About the Mean. As shown in Section 2 above, moments about the mean may be defined as follows:

$$\sum_{s=0}^{\infty} \mu_s \frac{\alpha^s}{s!} = \sum_{x=0}^{\infty} A_x e^{-m_1\alpha} (e^\alpha - 1)^x \tag{4.1}$$

where m_1 is the first moment about the origin:

$$\begin{aligned} m_1 &= np \quad \textit{Binomial} \\ &= a \quad \textit{Poisson} \\ &= lr/n \quad \textit{Hypergeometric} \end{aligned}$$

Now

$$\begin{aligned} \sum_{x=0}^{\infty} A_x e^{-m_1\alpha} (e^\alpha - 1)^x &= \sum_{x=0}^{\infty} A_x \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} e^{(v-m_1)\alpha} \\ &= \sum_{s=0}^{\infty} \frac{\alpha^s}{s!} \sum_{x=0}^{s} x!\, A_x\, \sigma_{x,s} \end{aligned}$$

where

$$x!\, \sigma_{x,s} = \sum_{v=0}^{x} (-1)^{x-v} \binom{x}{v} (v - m_1)^s = \Delta^x (-m_1)^s.$$

It will be observed that for $m_1 = 0$, $\sigma_{x,s} = S_{x,s}$. The internal series terminates at s for the same reason as before.

The moments about the mean are then given by:

$$\mu_s = \sum_{x=0}^{s} x!\, A_x\, \sigma_{x,s} \tag{4.2}$$

The particular forms for the three distributions are as follows:

$$\mu_s = \sum_{x=0}^{s} (n)_x p_x \sigma_{x,s} \qquad \textit{Binomial} \tag{4.3}$$

$$\mu_s = \sum_{x=0}^{s} a^x \sigma_{x,s} \qquad \textit{Poisson} \tag{4.4}$$

$$\mu_s = \sum_{x=0}^{s} \frac{(l)_x (r)_x}{(n)_x} \sigma_{x,s} \qquad \textit{Hypergeometric.} \tag{4.5}$$

The coefficients $\sigma_{x,s}$ satisfy the following recurrence relation:[4]

$$\sigma_{x,s+1} = (x - m_1)\sigma_{x,s} + \sigma_{x-1,s} \tag{4.6}$$

which in conjunction with equations (4.3)–(4.5) leads to moment recurrence relations as before. The actual derivation is somewhat complicated by the circumstance that $\sigma_{x,s}$ is a function of m_1 and therefore of the frequency parameters, rather than a constant as before. The derivation is illustrated for the binomial distribution as follows:

$$\begin{aligned}
\mu_{s+1} &= \sum_{x=0}^{s+1} (n)_x p^x \sigma_{x,s+1} \\
&= \sum_{x=0}^{s+1} (n)_x p^x [(x - np)\sigma_{x,s} + \sigma_{x-1,s}] \\
&= \sum_{x=0}^{s} (n)_x \sigma_{x,s}\, pD_p(p^x) - np\mu_s + \sum_{x=0}^{s+1} (n)_x p^x \sigma_{x-1,s} \\
&= pD_p\mu_s + nsp\mu_{s-1} - np\mu_s + np\mu_s - p^2[D_p\mu_s + ns\mu_{s-1}] \\
&= pq\,[ns\mu_{s-1} + D_p\mu_s].
\end{aligned}$$

The steps in the process are expanded as follows:

$$\begin{aligned}
\sum_{x=0}^{s} (n)_x \sigma_{x,s}\, pD_p(p^x) &= \sum^{s} (n)_x [pD_p(p^x \sigma_{x,s}) - p^x\, pD_p(\sigma_{x,s})] \\
&= pD_p\mu_s - p \sum_{x=0}^{s} (n)_x p^x(- ns\sigma_{x,s-1}) \\
&= pD_p\mu_s + nsp\mu_{s-1}
\end{aligned}$$

$$\begin{aligned}
\sum_{x=0}^{s+1} (n)_x p^x \sigma_{x-1,s} &= \sum_{x=0}^{s+1} (n - x + 1)\,(n)_{x-1} p^x \sigma_{x-1,s} \\
&= n \sum^{s} (n)_x p^{x+1} \sigma_{x,s} - \sum_{x=0}^{s} x(n)_x p^{x+1} \sigma_{x,s} \\
&= np\,\mu_s - p^2\,[D_p\,\mu_s + ns\,\mu_{s-1}].
\end{aligned}$$

The relation $D_p\sigma_{x,s} = -ns\sigma_{x,s-1}$ is obtained from the definition equation of $\sigma_{x,s}$ (with $m_1 = np$).

The resulting recurrence relations for the three distributions are as follows:

$$\mu_{s+1} = nspq\,\mu_{s-1} + pq\,D_p\,\mu_s \qquad \textit{Binomial} \tag{4.7}$$

$$\mu_{s+1} = as\mu_{s-1} + a\,D_a\,\mu_s \qquad \textit{Poisson} \tag{4.8}$$

[4] Jordan, loc. cit. or E. C. Molina, *An Expansion for Laplacian Integrals . . .*, Bell System Technical Journal, **11**, p. 571.

$$(4.9)\qquad \mu_{s+1} = (n+1)\left[\mu_s - \sum_{v=0}^{s}\binom{s}{v}K_1^v\,\mu_{s-v}(l, r, n+1)\right] \quad \textit{Hypergeometric}$$
$$-\frac{lr}{n}\left[\mu_s - \sum_{v=0}^{s}\binom{s}{v}K_2^v\,\mu_{s-v}(l-1, r-1, n-1)\right]$$

where

$$K_1 = \frac{-lr}{n(n+1)} = \Delta_n \frac{lr}{n}$$
$$K_2 = \frac{(l-1)(r-1)}{(n-1)} - \frac{lr}{n}.$$

The last of these, which appears to be new, seems to be of formal interest only.

The coefficients $\sigma_{x,s}$ are related to the Stirling numbers by the expression:

$$\sigma_{x,s} = \sum_{v=0}^{s-x}(-1)^v\binom{s}{v}S_{x,s-v}\,m_1^v = \sum_{v=0}^{s-x}a_v m_1^v$$

and consequently can be exhibited with detached coefficients in the form $a_0 + a_1 + a_2 + \cdots + a_{s-x}$. For the binomial and Poisson distributions certain simplifications, to be developed in the section following, in equations (4.3) and (4.4) may be made. For the hypergeometric distribution it appears necessary to use equation (4.5); the following short table of $\sigma_{x,s}$, employing the detached coefficients mentioned above, is given for this purpose:

1	0−1	1				
2	0+0+1	1−2	1			
3	0+0+0−1	1−3+3	3−3	1		
4	0+0+0+0+1	1−4+6−4	7−12+6	6−4	1	
5	0+0+0+0+0−1	1−5+10−10+5	15−35+30−10	25−30+10	10−5	1

5. **Binomial and Poisson Moments About the Mean—Simplified Formulas.**

5.1 **Binomial.** From examination of the first few moments about the mean, it appears expedient[5] to write the formulas:

$$(5.1.1)\qquad \begin{aligned}\mu_{2s} &= \sum_{x=1}^{s}\alpha_{x,2s}(npq)^x\\ \mu_{2s+1} &= (q-p)\sum^{s}\alpha_{x,2s+1}(npq)^x\end{aligned}$$

[5] The kind of expression chosen admits of some variety. A recurrence relation for coefficients in the expansion $\mu_s = \sum_{x=1}^{s}\alpha_{x,s}\,p^x$ has been given by E. H. Larguier, *On a Method For Evaluating the Moments of a Bernoulli Distribution*, Bull. Am. Math. Soc., **42**, 1, p. 24 (Abstract 8); I am indebted to Mr. Larguier for the opportunity of examining his results in advance of publication.

When these are substituted into the moment recurrence relation, the coefficients are found to be related as follows:

$$\alpha_{x,2s} = [x + pqD_{pq}]\alpha_{x,2s-1} + (2s - 1)\alpha_{x-1,2s-2}$$
$$-2pq[1 + 2x + 2pqD_{pq}]\alpha_{x,2s-1}$$
$$\alpha_{x,2s+1} = [x + pqD_{pq}]\alpha_{x,2s} + 2s\alpha_{x-1,2s-2}$$

or, in general,

$$\begin{aligned} \alpha_{x,s+1} &= [x + pqD_{pq}]\alpha_{x,s} + s\alpha_{x-1,s-1} \\ &\quad - pq[1 - (-1)^s]\,[1 + 2x + 2pqD_{pq}]\alpha_{x,s} \end{aligned} \tag{5.1.2}$$

Using detached coefficients of powers of pq as outlined above, these coefficients may be exhibited as follows:

$s \backslash x$	$\alpha_{x,s}$ 1	2	3	4
2	1			
3	1			
4	1 − 6	3		
5	1 − 12	10		
6	1 − 30 + 120	25 − 130	15	
7	1 − 60 + 360	56 − 462	105	
8	1 − 126 + 1680 − 5040	119 − 2156 + 7308	490 − 2380	105
9	1 − 252 + 5040 − 20160	246 − 6948 + 32112	1918 − 13216	1260

It may be noted that the coefficients of the first column in conjunction with equations (5.1.1) give the binomial seminvariants.

Equations (5.1.1) make the coefficients functions of pq only; a slight alteration makes the coefficients functions of n only. Thus:

$$\begin{aligned} \mu_{2s} &= \sum_{x=1}^{s} \beta_{x,2s}\,(pq)^x \\ \mu_{2s+1} &= (q - p)\sum_{x=1}^{s} \beta_{x,2s+1}\,(pq)^x \end{aligned} \tag{5.1.3}$$

and the coefficients are found to satisfy the recurrence relation:

$$\beta_{x,s+1} = x\beta_{x,s} + ns\beta_{x-1,s-1} - [1 - (-1)^s](2x - 1)\beta_{x-1,s}. \tag{5.1.4}$$

These coefficients may be exhibited by a rearrangement of the table given

above as may be seen by comparing equations (5.1.1) and (5.1.3). The first few coefficients are as follows:

$s \backslash x$	$n^{-1}\beta_{s,s}$ 1	2	3
2	1		
3	1		
4	1	$-6+3$	
5	1	$-12+10$	
6	1	$-30+25$	$120-130+15$

5.2 Poisson. The Poisson moments about the mean may be expressed as follows:

$$\mu_s = \sum_{x=0}^{[s/2]} \alpha_{x,s}\alpha^x \tag{5.2.1}$$

where [ ] represents "integral part of" and

$$\alpha_{x,s+1} = x\alpha_{x,s} + s\alpha_{x-1,s-1}. \tag{5.2.2}$$

The coefficients $\alpha_{x,s}$ are the constant terms in the expressions for the corresponding binomial distribution coefficients in powers of pq.

BELL TELEPHONE LABORATORIES.

NOTE ON ZOCH'S PAPER ON THE POSTULATE OF THE ARITHMETIC MEAN

By Albert Wertheimer

1. **Introduction.** There appeared recently a paper by Richmond T. Zoch[1] entitled "On The Postulate of the Arithmetic Mean." The stated purpose of his paper, was to show that the derivation of the Postulate as given by Whittaker & Robinson, is not correct. It is the purpose of this paper to show, that Zoch has not proven any error to exist in the Whittaker & Robinson derivation, but that there are a few errors in his paper. As this paper is intended to be read with Zoch's paper as a reference, the terms used there will not be redefined here, and except where otherwise stated, the symbols used will have the same meaning.

2. Zoch introduces the function

$$f \equiv \bar{x} + a\mu_3/\mu_2$$

and claims that it satisfies all the four axioms of Whittaker & Robinson, and obviously it is not the arithmetic mean. He therefore concludes that their derivation must have errors somewhere, and proceeds to find them. Let us first examine the f function. Considering only the part μ_3/μ_2, the partial derivatives with respect to x_i are given by

$$\frac{3\mu_2\{(x_i - \bar{x})^2 - \mu_2\} - 2\mu_3(x_i - \bar{x})}{n\mu_2^2}$$

It is then stated (p. 172) ". . . clearly these partial derivatives are single valued and continuous. Therefore the function μ_3/μ_2 satisfies axiom IV." Now, the condition that a function be continuous and single valued means of course that this be true throughout the region of definition of the function. It is not shown how these derivatives are clearly continuous and single valued for the very important case where all the x's are equal and the derivatives become indeterminate. As a matter of fact they are not continuous in this case, and therefore the f function does not satisfy axiom IV. To prove this, we only have to consider the very simple case where we let

$$x_i = k + c_i z$$

[1] This Journal Vol. VI no. 4, Dec. 1935, pp. 171–182.

where k is a fixed constant, c_i is a set of arbitrary constants not all equal, and z is a parameter. We then have

$$\bar{x} = k + \bar{c}z$$
$$\mu_2 = \mu_2' z^2$$
$$\mu_3 = \mu_3' z^3$$

where

$$\bar{c} = 1/n \sum c_i$$
$$\mu_2' = 1/n \sum (c_i - \bar{c})^2$$
$$\mu_3' = 1/n \sum (c_i - \bar{c})^3$$

Substituting these values in f and the derivatives, we get taking $a = 1$,

$$f = k + z\bar{c} + z^3\mu_3'/z^2\mu_2'$$

$$\partial f/\partial x_i = 1/n + \frac{3z^2\mu_2'\{z^2(c_i - \bar{c})^2 - z^2\mu_2'\} - 2z^4\mu_3'(c_i - \bar{c})}{nz^4\mu_2'^2}$$

Now going to the limit when z approaches zero, and all the x's approach k, we get

$$\lim_{z \to 0} f = k,$$

$$\lim_{z \to 0} \partial f/\partial x_i = 1/n\{-2 + 3(c_i - \bar{c})^2/\mu_2' - 2\,\mu_3'(c_i - \bar{c})/\mu_2'^2\}$$

Thus, when all the x's approach the same value, the function f also approaches the same value independent of the c's, that is regardless of the mode of approach, while the derivatives can take on any value depending on the c's that is on how the limiting value of f is approached. The f function then does not have continuous single valued partial derivatives, and therefore does not satisfy axiom IV.

In part 2 of the paper it is stated "Now when the x_i all approach a then both f and $\partial f/\partial x_j$ become indeterminate forms. However, in this case f takes an indeterminate form which can be evaluated and it can be shown that μ_3/μ_2 will always have the value zero, i.e., f will have the value a when all the $x_i \to a$; while the $\partial f/\partial x_i$ can take any value whatever and in general the $\partial f/\partial x_i$ will not be equal when the $x_i \to a$." This statement really amounts to saying that the f function does not satisfy axiom IV, but it is there used to demonstrate that one of Schiaparelli's propositions is false.

3. Having exhibited a function different from the arithmetic mean, and supposedly satisfying all the four axioms, the question is asked "Where is the proof given by Whittaker & Robinson lacking in rigor?" After numbering the various steps in the derivation "... for the sake of rigor and careful reasoning

. . ." it is stated (p. 174), "The sixth step involves the tacit assumption that the partial derivatives are functions of k. These partial derivatives are not necessarily functions of k . . ." and it is therefore concluded that the sixth step is not valid. Now, how can any function that by definition is to be evaluated at $\theta k x_i$ not be a function of k? What is shown (pp. 174–5) is that these derivatives do not necessarily involve k explicitly, but this is neither implied nor necessary for the sixth step, and there is no ground for doubting its validity.

4. In order to overcome the supposed defect in the sixth step, it is proposed to change axiom IV so as to require the partial derivatives to be constants. But even then (p. 175) ". . . there remains an objection in the seventh step." Now, the seventh step consists of the statement that if

$$\phi(x_i) = \sum c_i x_i$$

where the c's are independent of the x's then due to the condition that ϕ be a symmetric function, all the c's must be equal. To show the defect in this step it is stated, that under certain conditions ". . . the function $f \equiv \bar{x} + \mu_3/\mu_2$ will have partial derivatives with respect to x_i which are unequal and constant; yet at the same time the function f is a symmetrical expression of the n variables." Granting that all that is correct, what has this got to do with the seventh step? The f function certainly is not of the type $\sum c_i x_i$ to which the seventh step is applied.

5. One more point should be mentioned. On p. 181 it is supposedly proven that any function satisfying the first three axioms must have continuous first partial derivatives. The proof is essentially as follows: Assuming all the x's are given the same increment Δx, the increment of the function then is $\Delta\phi$. It is then stated ". . . but by axiom I, $\Delta\phi = \Delta x$. Therefore $\Delta\phi/\Delta x = 1 = d\phi/dx$. In other words, the total derivative of ϕ exists and is constant. Therefore the total derivative of ϕ is continuous." From this, the continuity of the first partial derivatives is proven by means of Euler's Theorem for homogeneous functions. Now, just what does the symbol $d\phi/dx$ (which is called the total derivative) mean for a function of many independent variables? Besides, (whatever this symbol means) is it considered rigorous to deduce a general Theorem from the very special case where all the differentials are made equal? This is one place where the f function could be used effectively as an exhibit of a function satisfying the first three axioms, and not having continuous partial derivatives.

It is also stated (p. 181) that ". . . it would seem more satisfactory to postulate that the function ϕ is single valued, for the single-valuedness of a derivative does not insure the single-valuedness of the integral while the single-valuedness of a function does insure the single-valuedness of the derivative where the derivative exists." This statement is certainly not self evident and requires

proof. For a single variable at least, it is easy to imagine a function represented by a curve with corners defined in a certain interval. The function then could be single valued everywhere in the interval, while the derivatives at the corners may exist and have two distinct values, depending on whether the corner is approached from the right or the left. On the other hand it is hard to imagine a curve representing a single valued function such that the integral i.e. the function represented by the area under the curve should not be single valued.

6. **In Conclusion:** It is stated in the Introduction that "Since this book has had wide circulation, it is believed that the errors in this proof should be called to the attention of the users of the book. The present paper has been prepared for this purpose." It is for the same reason, that this paper was prepared to show that no error has been proven to exist.

Bureau of Ordnance, U. S. Navy Department

NOTE ON THE BINOMIAL DISTRIBUTION

By C. E. Clark

The purpose of this note is to show that

$$f(x) = (-1)^n \frac{q^n n!}{\pi} \left(\frac{p}{q}\right)^x \frac{\sin \pi x}{x^{(n+1)}} \tag{1}$$

where n is an integer $\geqq 0$, $0 < p < 1$, $p + q = 1$, and $x^{(n+1)} = x(x-1)(x-2) \cdots (x-n)$, is a function whose values at $x = 0, 1, 2, \cdots n$ are the successive terms of the expansion of $(q + p)^n$, and also to consider the problem of fitting $f(x)$ to an observed frequency distribution.

The statement made about (1) can be verified by evaluating (1) as an indeterminate form. On the other hand, (1) can be derived by observing that the x-th term (x an integer) of the expansion of $(q + p)^n$ is

$$\frac{n!}{x!(n-x)!} p^x q^{n-x} = \frac{\Gamma(n+1) p^x q^{n-x}}{\Gamma(x+1)\Gamma(n-x+1)}; \tag{2}$$

then (1) can be derived from (2) by means of the product expansions for $\Gamma(x)$ and $\sin x$. This derivation of (1) from (2) can also be carried out by expressing (2) as a Beta function and then using

$$B(x+1, n-x+1) = \int_0^\infty \frac{t^x}{(1+t)^{n+2}}\, dt = (-1)^n \frac{\pi}{(n+1)!} \frac{x^{(n+1)}}{\sin \pi x}.$$

This integration can be performed by means of the theory of residues.

Consider the problem of fitting (1) to an observed frequency distribution. We shall write (1) in the form

$$F(z) = ab^x \frac{\sin \pi x}{x^{(n+1)}}, \qquad x = \frac{nb}{1+b} + h(z - \bar{z}) \tag{3}$$

and determine the constants a, b, n, and h so that, when $\bar{z}$ is the mean of the observed distribution, $F(z)$ will fit the distribution.

The values of a, b, n, and h can be determined by the method of moments. Let ν_2, ν_3, and ν_4, denote the usual second, third, and fourth moments of the distribution, which are calculated in the usual way (as in W. P. Elderton, *Frequency-Curves and Correlation*) and not adjusted by any procedure such as Sheppard's adjustments. Also, use the usual notation $\beta_1 = \frac{\nu_3^2}{\nu_2^3}$ and $\beta_2 = \frac{\nu_4}{\nu_2^2}$.

Then, the method of moments gives

$$n = \frac{2}{3 + \beta_1 - \beta_2} \tag{4}$$

$$b = \frac{2 + n\beta_1 \pm \sqrt{n\beta_1(4 + n\beta_1)}}{2} \tag{5}$$

$$h = \sqrt{\frac{nb}{\nu_2}}\left(\frac{1}{1+b}\right)$$

$a = (-1)^n \dfrac{h(\Sigma f)n!}{\pi(1+b)^n}$, where Σf is the sum of the frequencies of the distribution. An integer n is chosen nearest the value assigned by (4). The two values of b from (5) determine two curves that are congruent but whose skewnesses are of opposite sign. Hence, b is uniquely determined by (5) and the sign of the skewness of the data.

For a symmetrical distribution, $b = 1$, $\nu_3 = 0$, and

$$n = \frac{2}{3 - \beta_2}$$

$$h = \frac{\sqrt{n}}{2\sqrt{\nu_2}}.$$

We shall consider an illustrative example. In the following table the columns $f(z)$ and $f_2(z)$ are taken from W. P. Elderton, *Frequency-Curves and Correlation* (1906), page 62. $f(z)$ is an empirical frequency distribution, while $f_2(z)$ is obtained by fitting a Pearson Type II curve to the distribution $f(z)$. $f_1(z)$ is computed from

$$f_1(z) = 1624\,\frac{\sin \pi x}{x^{(6)}}, \qquad x = 2.0973 + .808z$$

which is determined by the method of this note. $f_3(z)$ is obtained by fitting the normal curve

$$f_3(z) = 485.1e^{-\frac{(z-.4985)^2}{2(1.829)}}.$$

z	$f(z)$	$f_1(z)$	$f_2(z)$	$f_3(z)$
−3	11	18	14	19
−2	116	107	109	92
−1	274	281	286	263
0	451	438	433	444
1	432	437	433	444
2	267	267	285	263
3	116	106	109	92
4	16	18	14	19

The coefficients of goodness of fit for $f_1(z)$, $f_2(z)$, and $f_3(z)$ are respectively .35, .58, and .02.

CONVEXITY PROPERTIES OF GENERALIZED MEAN VALUE FUNCTIONS[1]

By Nilan Norris

Consider the following generalized mean value functions: (1) the unit weight or simple sample form, $\phi(t) = \left(\frac{x_1^t + x_2^t + \cdots + x_n^t}{n}\right)^{\frac{1}{t}}$, in which the x_i are positive real numbers not all equal each to each, and in which t may take any real value; (2) the weighted sample form, $\omega(t) = \left(\frac{c_1x_1^t + c_2x_2^t + \cdots + c_nx_n^t}{c_1 + c_2 + \cdots + c_n}\right)^{\frac{1}{t}}$, in which the c_i are positive numbers not all equal each to each, and in which the x_i and t are restricted as in $\phi(t)$; (3) the integral form, $\theta(t) = \left[\int_{x=0}^{1} x^t dx\right]^{\frac{1}{t}}$, where $\int_{x=0}^{1} x^t dx$ exists for every real value of t; and (4) the generalized integral form $\Phi(t) = \left[\int_{x=0}^{\infty} x^t d\psi(x)\right]^{\frac{1}{t}}$, where $\psi(x)$ is a non-decreasing function integrable in the Riemann-Stieltjes sense such that $\psi(\infty) - \psi(0) = 1$, and such that $\int_{x=0}^{\infty} x^t d\psi(x)$ exists for every real value of t. The facts that all of these functions are monotonic increasing and that both $\phi(t)$ and $\omega(t)$ have two horizontal asymptotes have been previously demonstrated.[2] Although the existence of $\phi(t)$ and $\omega(t)$ has been known since 1840, there appears to have been no attempt made to investigate the behavior of the second derivatives of them.[3]

When the x_i are price relatives, production relatives, or similar data, $\phi(t)$ and $\omega(t)$ yield common types of index numbers by direct substitution of integral values of t. For any values of t such that $0 < t_1 < t_2 < \infty$, the type bias of $\phi(t_2)$ will be greater than the type bias of $\phi(t_1)$. Similarly, for any values of t such that $-\infty < t_1 < t_2 < 0$, the type bias of $\phi(t_1)$ will be greater than the type bias of $\phi(t_2)$. The second derivatives of $\phi(t)$ and $\omega(t)$ indicate whether

[1] Presented at a joint meeting of the American Mathematical Society, the Econometric Society, and the Institute of Mathematical Statistics at St. Louis on January 2, 1936. The writer is indebted to C. C. Craig, Einar Hille, Dunham Jackson, and J. Shohat for helpful critical reviews of the preliminary draft of this paper.

[2] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge University Press, London, 1934), pp. 12–15; and Nilan Norris, "Inequalities among Averages," *Annals of Mathematical Statistics*, Vol. VI, No. 1, March, 1935, pp. 27–29.

[3] Jules Bienaymé, *Société Philomatique de Paris*, Extraits des procès-verbaux des séances pendant l'année 1840 (Imprimerie D'A. René et Cie., Paris, 1841), Séance du 13 juin 1840 p. 68.

type bias is changing at an increasing or a decreasing rate as between the unlimited number of averages available for use. Considerable interest attaches to $\omega(t)$, the weighted sample form of function.

Let $\omega(t)$ be made arbitrary for the case of $n = 2$, with $x_1 = 1$, and $x_2 = e^{-\lambda}$, where λ is any real number. Also let $c_1 = \alpha$, and $c_2 = \beta$, where $\alpha + \beta = 1$. Then $\omega(t) = [\alpha + \beta e^{-\lambda t}]^{\frac{1}{t}}$. Now for all values of t,

$$\alpha + \beta e^{-\lambda t} = 1 - \frac{\beta\lambda}{1} t + \frac{\beta\lambda^2}{2} t^2 - \frac{\beta\lambda^3}{6} t^3 + \cdots$$

For $|t|$ sufficiently small, it follows that

$$\log(\alpha + \beta e^{-\lambda t}) = -\beta\lambda t + \tfrac{1}{2}\beta\lambda^2(1-\beta)t^2 + \beta\lambda^3\left[-\frac{1}{6} + \frac{\beta}{2} - \frac{\beta^2}{3}\right]t^3 +$$

so that for $t \neq 0$

$$\frac{1}{t}\log(\alpha + \beta e^{-\lambda t}) = -\beta\lambda + \tfrac{1}{2}\beta\lambda^2(1-\beta)t + \beta\lambda^3\left[-\frac{1}{6} + \frac{\beta}{2} - \frac{\beta^2}{3}\right]t^2 +$$

Therefore $\omega(t) = \exp.\left[\frac{1}{t}\log(\alpha + \beta e^{-\lambda t})\right]$

$$= e^{-\beta\lambda}\left[1 + \tfrac{1}{2}\beta\lambda^2(1-\beta)t + \beta\lambda^3\left\{-\frac{1}{6} + \frac{\beta}{2} - \frac{\beta^2}{3} + \frac{1}{8}\beta(1-\beta)^2\lambda\right\}t^2 + \cdots\right].$$

It follows that $\omega''(0) = 2\beta\lambda^3 e^{-\beta\lambda}\left[-\frac{1}{6} + \frac{\beta}{2} - \frac{\beta^2}{3} + \frac{1}{8}\beta(1-\beta)^2\lambda\right]$. It is clear that $\omega(0)$ is the weighted geometric mean, and that $\phi(0)$ is the unit weight or simple sample form of geometric mean. As a means of demonstrating the range of values which $\omega''(0)$ may take it is helpful to rewrite the expression for $\omega''(0)$ as follows:

$$\omega''(0) = \tfrac{1}{4}\beta^2(1-\beta)^2\lambda^3\left[\lambda - \frac{4}{3}\frac{1-2\beta}{\beta(1-\beta)}\right]e^{-\beta\lambda} \equiv f(\lambda, \beta).$$

This consideration makes it possible to distinguish three cases of $y = f(\lambda, \beta)$ for fixed β, namely, $0 < \beta < \frac{1}{2}$; $\beta = \frac{1}{2}$; and $\frac{1}{2} < \beta < 1$. In all three cases $f(\lambda, \beta)$ has an absolute minimum $\mu(\beta) \leqq 0$, and $\mu(\frac{1}{2}) = 0$. The corresponding values of λ satisfies the quadratic equation $\lambda^2 - \frac{4}{3}\frac{4-5\beta}{\beta(1-\beta)}\lambda + \frac{4-8\beta}{\beta^2(1-\beta)} = 0$. It is clear that by taking β near enough to 0, one can make $\mu(\beta)$ as large negative as is desired. Also, by choosing λ properly, one can make $\omega''(0)$ take any value between $\mu(\beta)$ and ∞. For example, when $\alpha = \beta = \frac{1}{2}$, λ may be selected so as to make $\omega''(0)$ any arbitrarily chosen non-negative number. For then $\omega''(0) = \frac{\lambda^4}{64}e^{-\frac{\lambda}{2}}$, and as λ increases from $-\infty$ to 0, $\omega''(0)$ decreases from ∞ to 0. If $\lambda = 0$, $\omega''(0) = 0$. If $\lambda > 0$, as λ increases from 0 to 8, $\omega''(0)$ increases to

$64e^{-4}$, and as λ increases beyond 8, $\omega''(0)$ decreases, approaching 0 as λ increases indefinitely. It is evident that the case of $\alpha = \beta = \frac{1}{2}$, with $\lambda = -\log 2$, $x_1 = 1$, and $x_2 = e^{-\lambda}$, is one in which $\omega(t)$ becomes the unit weight or simple sample type of generalized mean value function, namely, $\phi(t) = \left(\frac{1^t + 2^t}{2}\right)^{\frac{1}{t}}$. Reference to the first expression above noted for $\omega''(0)$ will make clear that $\phi''(0) = \frac{(\log 2)^4}{64}\sqrt{2}$ in this special case.

Analysis of $\Phi(t)$, the generalized integral form of generalized mean value function, makes it possible to characterize populations of a very general character, as well as samples. But in the case of $\Phi(t)$ it is even more difficult to generalize as to convexity properties. For example, let

$$\Phi(t) = \left[\int_{u=-\infty}^{\infty} e^{-ut}\, dE(u)\right]^{\frac{1}{t}},$$

where

$$E(u) = \frac{1}{\sqrt{\pi}} \int_{v=-\infty}^{u} e^{-v^2}\, dv.$$

This expression is obviously of the required generalized integral type. Now

$$[\Phi(t)]^t = \frac{1}{\sqrt{\pi}} \int_{u=-\infty}^{\infty} e^{-ut-u^2}\, du = \frac{1}{\sqrt{\pi}}\, e^{\frac{t^2}{4}} \int_{u=-\infty}^{\infty} e^{-\left(u+\frac{t}{2}\right)^2}\, du = e^{\frac{t^2}{4}}.$$

Therefore $\Phi(t) = e^{\frac{t}{4}}$, and $\Phi''(t) = \frac{e^{\frac{t}{4}}}{16} > 0$ for all t. That is, in this particular case, $\Phi(t)$ has only one horizontal asymptote.

The foregoing examples indicate that the following conclusions may be drawn as to the diverse convexity attributes of the various means as functions of t: (1) The unit weight form, $\phi(t)$, and the weighted sample form, $\omega(t)$, must always have a point of inflection, since both of them not only increase with t, but are doubly asymptotic (have two horizontal asymptotes). (2) Points of inflection for $\phi(t)$ and $\omega(t)$ do not necessarily occur at $t = 0$. (3) The generalized integral form, $\Phi(t)$, need not always have a point of inflection. That is, the second derivatives of certain forms of $\Phi(t)$ do not change their sign, since such forms are concave upward.

University of Michigan.

A SIMPLE FORM OF PERIODOGRAM

By Dinsmore Alter

Schuster's introduction of a method of systematic search for hidden periodicities and cycles opened a new field for the investigator of statistical data. The beauty of his method in its analogy to analysis of light, and the great reputation of its author, combined to give it universal acceptance and to blind statisticians to its faults.

In more recent years at least three new mathematical and two mechanical forms of periodogram analysis have been proposed, each of which exhibits certain advantages over the original one. The use of the term *periodogram* for these forms is an extension of Schuster's original definition which used as abscissae quantities proportional to the squares of the amplitudes of the sine terms found in the data for the various trial periods. He wrote: "It is convenient to have a word for some representation of a variable quantity which shall correspond to the spectrum of a luminous radiation. I propose the word *periodogram* and define it more particularly in the following way:

Let $\frac{1}{2}Ta = \int_{t_1}^{t_1+T} f(t) \cos ktdt$ and $\frac{1}{2}Tb = \int_{t_1}^{t_1+T} f(t) \sin ktdt$

where T may for convenience be chosen equal to some integer multiple of $\frac{2\pi}{k}$, and plot a curve with $\frac{2\pi}{k}$ as abscissae and $r = \sqrt{a^2 + b^2}$ as ordinates; this curve, or better, the space between this curve and the axis of abscissae, represents the periodogram of $f(t)$."

The following appear to be the essential criteria for a satisfactory form of periodogram:

1. It must exhibit plainly any repetition of form in the data regardless of how irregular the shape of the repeated interval may be. In doing this it must exaggerate the amplitude of the main terms at the expense of the lesser ones.
2. The calculation of the indices must be short. In a periodogram from many data the indices sometimes are computed for several hundred trial periods.
3. There should be a geometrical interpretation of the index used.
4. The frequency distribution of the index must be known.
5. Combining or smoothing the data should modify the index in a manner which leaves an obvious interpretation.

The Schuster periodogram has the following disadvantages:

1. Only sine terms of large amplitude are exhibited. A perfect repetition of an extremely irregular form of data would not be indicated in any way.

2. The calculations are long.

3. There is a considerable uncertainty in the length of the period found. Those methods of analysis which use harmonics as well as the fundamental have much less of this uncertainty.

The correlation periodogram has advantages in each of these points over the Schuster. However, even with it the calculations are fairly long. Furthermore, the modification of the coefficient introduced by grouping or smoothing is not a linear one.

The periodogram described here is a slight modification of one for which a preliminary note was published in 1933. Additional features have been studied and its applications to many data have shown its ease of calculation. This calculation has been reduced still more by a mechanical method which renders it practicable to contemplate the possibility of studying many data hitherto prohibited by excessive cost.

Consider data x_0, x_1, x_2, $\cdots x_i$, $\cdots x_{(n-1)}$. Let l be any integer less than n. Form the sum of the absolute values of $x_i - x_{(i-l)}$, designated by $\sum |x_i - x_{(i-l)}|$. Define $A = \sum_{i=l}^{n-1} \frac{|x_i - x_{(i-l)}|}{n-l}$, l takes the values of the various trial periods and is called the *lag*. A, therefore, is the mean error between prediction that data will be repeated after a lag of l and the fulfillment of the prediction. Such an index has a meaning that is immediately of use to a meteorologist or other investigator. Coefficients such as the Schuster and the correlation coefficient, although valuable statistically, are of less immediate interest.

The standard deviation of these errors of prediction follows at once from standard formulae under assumption of normal distribution.

$$\sigma = 1.25\,A$$

The distribution of σ, as computed from the absolute values of data, has been studied by Helmert and by Fisher. Davies and E. S. Pearson have compared the various methods of estimating σ. For the large number, $(n - l)$, pairs of data used for a periodogram point, this method becomes almost as precise as the usual one which would square the values of $(x_i - x_{i-l})$. For $(n - l)$ as small as 50, the standard deviation of the standard deviation by this method is only seven percent larger than by the other one. Fisher has shown that

$$\sigma_\sigma \to \frac{\sigma}{\sqrt{n-l}} \sqrt{\frac{\pi - 2}{2}} \qquad \text{as } (n - l) \to \infty$$

This may be written as

$$\sigma_\sigma \to \frac{1.068\,\sigma}{\sqrt{2(n-l)}}$$

The distribution approaches normal rapidly and for all values of $(n - l)$ that would be used in periodogram calculation certainly may be considered as normal. It will be very seldom that a value of $(n - l)$ much smaller than 200 will be used.

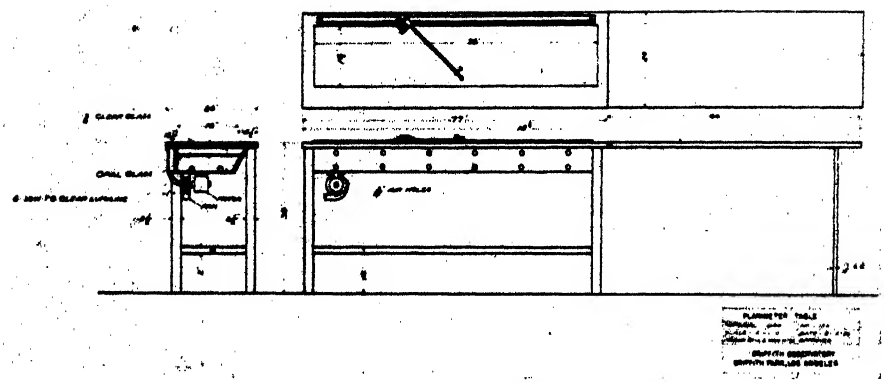
The data may be printed on two strips of adding machine tape held together by clips so as to match data separated by a lag l. In arranging them for investigation, it usually is most convenient to make all numbers positive. The computer subtracts mentally and puts the difference into an adding machine, which gives him A almost immediately.

For some computers, and especially where the numbers are large, another method of obtaining A may save time or lead to less numerical mistakes. The computer will form the sum of all his data. He will, as for the other form of computation, put these on two pieces of adding machine tape that he lays side by side. However, instead of putting the difference of the pairs into the machine, he will, in each case, put in the smaller datum of the pair. Then,

$$(n - l)A_l = 2 \sum \text{all data} - [\sum \text{1st } (n - l) + \sum \text{last } (n - l) \text{ data}] - 2 \sum \text{smaller}$$

The derivation of this equation is obvious. In computing by this method the subtotaler on the machine can be used to make the strip of sums of the first $(n - l)$ data and of the last $(n - l)$ for all values of l. The first term on the right hand side is a constant, the last is twice the sum of the smaller numbers chosen in the pairs. I have computed by both methods, and where the numbers are small, I prefer the former. Where they are large, I prefer the latter. However, when one must use comparatively untrained computers, he will find less mistakes made if the computer does not make the subtractions.

The calculation of A is much shorter than that for the indices even of the correlation and variance periodograms. It may, however, be shortened even more by a mechanical arrangement. $(n - l)A_l$ is the area between two histograms of the data matched after a lag l. These may be carefully graphed on a large scale and two such graphs superposed over a table with a translucent illuminated top. On the edge of this table is the track to guide a rolling planimeter. A, as computed by this means, is accurate to approximately one-half of one percent of its value, a much more exact value than is needed. The details of such a device as constructed for the Griffith Observatory are shown by the accompanying photograph and diagram. The dual saving of time by the method and by its mechanical application have resulted in the adoption of a much more ambitious program of meteorological research than previously was contemplated.



SCALE DIAGRAM OF PLANIMETER DEVICE

PLANIMETER DEVICE FOR MECHANICAL CALCULATION

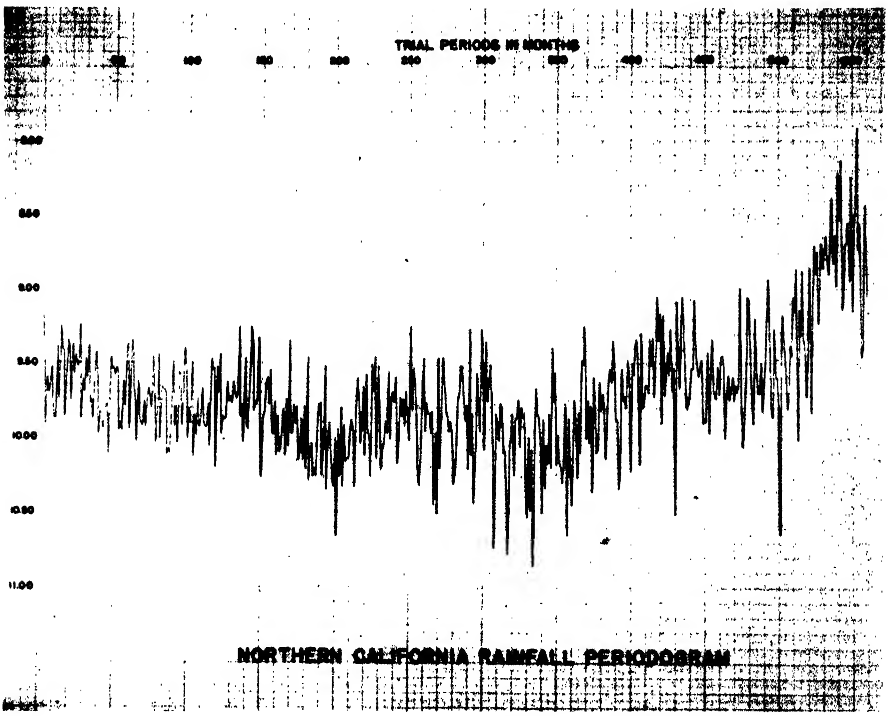
The form taken by the periodogram is important. Consider the simplest case, data which follow a sine curve.

$$y_i = a \cos\left(\frac{2\pi i - c}{p}\right)$$

$$y_i - y_{i-l} = 2a \sin\frac{\pi l}{p}\left[\sin\frac{2\pi(\frac{1}{2}l - i) + c}{p}\right]$$

The term in brackets takes values distributed around the circle and the part outside is a constant for any one lag. The bracket term sums approximately to $\dfrac{2(n-l)}{\pi}$, since we consider all terms as of one sign only.

$$\therefore A_l = \left|\frac{4a}{\pi}\sin\frac{\pi l}{p}\right|$$



If the absolute values were not considered in the expression for A_l, the periodogram would be a sine curve of period $2p$. The lack of sign gives a cusp curve with the cusp at lags p, $2p$, etc. Such a form is advantageous in that the periodogram gives sharp peaks at multiples of the periods which may exist.

The effect of the periodogram in exaggerating the principal terms at the expense of the smaller ones may be obtained most easily by equating σ as obtained by the linear and the quadratic formulae.

The data may be written as the sum of cosine terms

$$y_i = a\cos\left(\frac{2\pi i - \varphi_a}{p_a}\right) + b\cos\left(\frac{2\pi i - \varphi_b}{p_b}\right) + \cdots + c_i$$

$$y_i - y_{i-l} = 2a\sin\frac{\pi l}{p_a}\left[\sin\frac{2\pi(\frac{1}{2}l - i) + \varphi_a}{p_a}\right] + \cdots + (c_i - c_{i-l})$$

$$\sum(y_i - y_{i-l})^2 = 2(n-l)a^2\sin^2\frac{\pi l}{p_a} + 2(n-l)b^2\sin^2\frac{\pi l}{p_b} + \cdots + (n-l)\sqrt{2}\,\sigma_c^2$$

The sine terms contribute to A_l^2 in proportion to the squares of their amplitudes. On account of the $\sin^2\dfrac{\pi l}{p_i}$ factor, they contribute very little to values of A_l for which $\dfrac{\pi l}{p_i}$ is not very closely an even multiple of π.

This method has been applied to rainfall data of the Pacific Coast and has proved as satisfactory in practice as would be expected from the simplicity of the theory. The periodogram of rainfall stations along the northern third of the California coast is shown here, exhibiting perhaps the most definite single piece of evidence ever found for rainfall cycles. Outstanding is a cycle of about 45 years with its fourth harmonic as the secondary feature. The writer expects to publish the results of that work in the Monthly Weather Review.

ON CERTAIN DISTRIBUTIONS DERIVED FROM THE MULTINOMIAL DISTRIBUTION[1]

By Solomon Kullback

1. Introduction. With the multinomial distribution as a background, there may be derived a number of distributions which are of interest in certain practical applications. Several of these distributions are here presented and the theory is illustrated by specific examples.

2. Preliminary data. In the discussion of the distributions to be considered there are needed certain factorial sums whose values are now to be derived. In the following discussion only positive integral values (including zero) are to be considered.

There is desired the value, in terms of N, n, r, of

$$f_r(n, N) = \sum \frac{N!}{x_1!\, x_2! \cdots x_n!} \tag{2.1}$$

where the summation is for all values of $x_1, x_2, \cdots, x_n$ such that $x_1 + x_2 + \cdots + x_n = N$ and no x is equal to r.

Let us first consider the case for $r = 0$; i.e., we desire a value for the sum in (2.1) for all values of $x_1, x_2, \cdots, x_n$ such that $x_1 + x_2 + \cdots + x_n = N$ and no x is equal to zero. By the multinomial theorem, we have that[2]

$$(a_1 + a_2 + \cdots + a_n)^N = \sum \frac{N!}{x_1!\, x_2! \cdots x_n!} a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n} \tag{2.2}$$

where the summation is for all values of $x_1, x_2, \cdots, x_n$ such that $x_1 + x_2 + \cdots + x_n = N$. If $a_1 = a_2 = \cdots = a_n = 1$, then

$$n^N = \sum \frac{N!}{x_1!\, x_2! \cdots x_n!}, \qquad x_1 + x_2 + \cdots + x_n = N. \tag{2.3}$$

The sum in (2.3) may however be rearranged into the sum of a number of terms as follows:

$$\begin{aligned}
&\sum\frac{N!}{x_1!\, x_2! \cdots x_n!}, && x_1 + x_2 + \cdots + x_n = N, && \text{no } x = 0;\\
&n \sum \frac{N!}{x_1!\, x_2! \cdots x_{n-1}!}, && x_1 + x_2 + \cdots + x_{n-1} = N, && \text{no } x = 0;\\
&\frac{n(n-1)}{2} \sum \frac{N!}{x_1!\, x_2! \cdots x_{n-2}!}, && x_1 + x_2 + \cdots + x_{n-2} = N, && \text{no } x = 0;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\binom{n}{r} \sum \frac{N!}{x_1!\, x_2! \cdots x_{n-r}!}, && x_1 + x_2 + \cdots + x_{n-r} = N, && \text{no } x = 0.
\end{aligned} \tag{2.4}$$

[1] Presented to the Institute of Mathematical Statistics January 2, 1936.

[2] H. S. Hall & S. R. Knight, *Higher Algebra*, MacMillan & Co., 4th Ed. (1924), Chap. 15.

Thus we may rewrite (2.3) as

$$n^N = f_0(n, N) + nf_0(n-1, N) \tag{2.5}$$
$$+ \frac{n(n-1)}{2!} f_0(n-2, N) + \cdots + \binom{n}{r} f_0(n-r, N) +$$

Replacing n by $n - 1$ in (2.5) there is obtained

$$(n-1)^N = f_0(n-1, N) \tag{2.6}$$
$$+ (n-1) f_0(n-2, N) + \cdots + \binom{n-1}{r} f_0(n-r-1, N) + \cdots$$

Multiplying (2.6) by n and subtracting the result from (2.5), there is obtained

$$n^N - n(n-1)^N = f_0(n, N) \tag{2.7}$$
$$- \frac{n(n-1)}{2!} f_0(n-2, N) - \cdots - r\binom{n}{r+1} f_0(n-r-1, N) - \cdots$$

Replacing n by $n - 2$ in (2.5) there is obtained

$$(n-2)^N = f_0(n-2, N) \tag{2.8}$$
$$+ (n-2) f_0(n-3, N) + \cdots + \binom{n-2}{r-1} f_0(n-r-1, N) + \cdots$$

Multiplying (2.8) by $n(n - 1)/2$ and adding the result to (2.8), there is obtained

$$n^N - n(n-1)^N + \frac{n(n-1)}{2!}(n-2)^N = f_0(n, N) + \frac{n(n-1)(n-2)}{3!} \tag{2.9}$$
$$f_0(n-3, N) + \cdots + \frac{r(r-1)}{2!}\binom{n}{r+1} f_0(n-r-1, N) + \cdots$$

Continuing this process, there is finally obtained the result that

$$f_0(n, N) = n^N - n(n-1)^N + \frac{n(n-1)}{2!}(n-2)^N - \cdots \pm n \cdot 1^N \tag{2.10}$$

It may be shown[3] that the right side of (2.10) is $\Delta^n x^N$ for $x = 0$. The author has elsewhere obtained (2.10), but by a special procedure not applicable to the general case.[4]

We may readily verify (2.10) for example, for $n = 3$, $N = 5$. If $x_1 + x_2 + x_3 = 5$ and no $x = 0$, then the sets of solutions are (3,1,1), (1,3,1), (1,1,3), (2,2,1), (2,1,2), (1,2,2), and $f_0(3,5) = 3 \cdot \frac{5!}{3!\,1!\,1!} + 3\frac{5!}{2!\,2!\,1!} = 150$. From (2.10) there is obtained $f_0(3,5) = 3^5 - 3.2^5 + 3.2/2 = 150$.

[3] E. T. Whittaker & G. Robinson, *The Calculus of Observations*, Blackie & Son Ltd. (1924), p. 7.

[4] S. Kullback, "On the Bernoulli Distribution," *Bull. Am. Math. Soc.*, December, 1935.

For the general case, we return again to (2.3) and rearrange the right side into the sum of a number of terms as follows:

$$
(2.11)\quad
\begin{aligned}
&\sum \frac{N!}{x_1!\,x_2!\cdots x_n!}, && x_1 + x_2 + \cdots + x_n = N, && \text{no } x = r;\\
&\frac{n}{r!}\sum \frac{N!}{x_1!\,x_2!\cdots x_{n-1}!}, && x_1 + x_2 + \cdots + x_{n-1} = N - r, && \text{no } x = r;\\
&\frac{n(n-1)}{2!\,(r!)^2}\sum \frac{N!}{x_1!\,x_2!\cdots x_{n-2}!}, && x_1 + x_2 + \cdots + x_{n-2} = N - 2r, && \text{no } x = r;\\
&\binom{n}{k}\left(\frac{1}{r!}\right)^k \sum \frac{N!}{x_1!\,x_2!\cdots x_{n-k}}, && x_1 + x_2 + \cdots + x_{n-k} = N - kr, && \text{no } x = r.
\end{aligned}
$$

Thus we may rewrite (2.3) as

$$
(2.12)\quad
\begin{aligned}
n^N &= f_r(n, N) + \frac{nN^{(r)}}{r!} f_r(n-1, N-r)\\
&\qquad + \frac{n(n-1)N^{(2r)}}{2!\,(r!)^2} f_r(n-2, N-2r) + \cdots
\end{aligned}
$$

where $N^{(k)} = N(N-1)(N-2)\cdots(N-k+1)$.

Replacing n by $n - 1$ and N by $N - r$ in (2.12) there is obtained

$$
(2.13)\quad
\begin{aligned}
(n-1)^{N-r} &= f_r(n-1, N-r)\\
&\qquad + \frac{(n-1)(N-r)^{(r)}}{r!} f_r(n-2, N-2r) + \cdots
\end{aligned}
$$

Multiplying (2.13) by $\dfrac{nN^{(r)}}{r!}$ and subtracting the result from (2.12), there is obtained

$$
(2.14)\quad n^N - \frac{nN^{(r)}}{r!}(n-1)^{N-r} = f_r(n, N) - \frac{n(n-1)N^{(2r)}}{2!\,(r!)^2} f_r(n-2, N-2r) - \cdots
$$

By continuing this process, in a manner similar to that used for the case $r = 0$ there is finally obtained

$$
(2.15)\quad
\begin{aligned}
f_r(n, N) &= n^N - \frac{nN^{(r)}}{r!}(n-1)^{N-r} + \frac{n(n-1)N^{(2r)}}{2!\,(r!)^2}(n-2)^{N-2r}\\
&\qquad - \binom{n}{3}\frac{N^{(3r)}}{(r!)^3}(n-3)^{N-3r} +
\end{aligned}
$$

By setting $r = 0$ in (2.15), there is of course obtained the value already found in (2.10).

We may readily verify (2.15) for example, for $n = 3$, $N = 5$, $r = 2$. If $x_1 + x_2 + x_3 = 5$ and no $x = 2$, then the sets of solutions are (5,0,0), (0,5,0),

(0,0,5), (4,1,0), (1,4,0), (1,0,4), (4,0,1), (0,1,4), (0,4,1), (3,1,1), (1,3,1), (1,1,3), and $f_2(3,5) = 3\cdot 5!/5! + 6\cdot 5!/4! + 3\cdot 5!/3! = 93$. From (2.15) there is obtained $f_2(3,5) = 3^5 - 3\cdot 5\cdot 4\cdot 2^3/2! + 3\cdot 2\cdot 5\cdot 4\cdot 3\cdot 2/2!(2!)^2 = 93$.

The same method of procedure may be applied to evaluate

$$f_{rs\cdots t}(n, N) = \sum \frac{N!}{x_1!\,x_2!\cdots x_n!}, \qquad x_1 + x_2 + \cdots + x_n = N, \quad \text{no } x = r, s, \cdots, \text{ or } t. \tag{2.16}$$

Thus, there is derived the result that

$$\begin{aligned} f_{rs}(n, N) &= n^N - n\left(\frac{N^{(r)}(n-1)^{N-r}}{r!} + \frac{N^{(s)}(n-1)^{N-s}}{s!}\right) \\ &\quad + n(n-1)\left(\frac{N^{(2r)}(n-2)^{N-2r}}{2!\,(r!)^2} + \frac{N^{(r+s)}(n-2)^{N-r-s}}{(r!)\,(s!)}\right. \\ &\quad \left. + \frac{N^{(2s)}(n-2)^{N-2s}}{2!\,(s!)^2}\right) - n(n-1)(n-2)\left(\frac{N^{(3r)}(n-3)^{N-3r}}{3!\,(r!)^3}\right. \\ &\quad \left. + \frac{N^{(2r+s)}(n-3)^{N-2r-s}}{2!\,(r!)^2\,(s!)} + \frac{N^{(r+2s)}(n-3)^{N-r-2s}}{2!\,(r!)\,(s!)^2} + \frac{N^{(3s)}(n-3)^{N-3s}}{3!\,(s!)^3}\right) \end{aligned} \tag{2.17}$$

We may readily verify (2.17) for example, for $n = 3$, $N = 5$, $r = 0$, $s = 2$. If $x_1 + x_2 + x_3 = 5$ and no $x = 0$ or 2, then the sets of solutions are (3,1,1), (1,3,1), (1,1,3) and $f_{02}(3,5) = 3\cdot 5!/3! = 60$. From (2.17) there is obtained $f_{02}(3,5) = 3^5 - 3(2^5 + 5\cdot 4\cdot 2^3/2) + 3\cdot 2(1/2! + 5\cdot 4/2! + 5\cdot 4\cdot 3\cdot 2/(2!)^3) = 60$. It will be shown later (see section 8) that

$$\begin{aligned} f_r(n, N) &= f_{rs}(n, N) + \frac{nN^{(s)}}{s!} f_{rs}(n-1, N-s) \\ &\quad + \frac{n(n-1)N^{(2s)}}{2!\,(s!)^2} f_{rs}(n-2, N-2s) + \cdots \end{aligned} \tag{2.18}$$

$$\begin{aligned} f_s(n, N) &= f_{rs}(n, N) + \frac{nN^{(r)}}{r!} f_{rs}(n-1, N-r) \\ &\quad + \frac{n(n-1)N^{(2r)}}{2!\,(r!)^2} f_{rs}(n-2, N-2r) + \cdots \end{aligned} \tag{2.19}$$

From (2.18) and (2.19) there may be derived, by a method similar to that employed in deriving (2.15), that

$$\begin{aligned} f_{rs}(n, N) &= f_r(n, N) - \frac{nN^{(s)}}{s!} f_r(n-1, N-s) \\ &\quad + \frac{n(n-1)N^{(2s)}}{2!\,(s!)^2} f_r(n-2, N-2s) - \cdots \end{aligned} \tag{2.20}$$

This latter result also follows from (2.17 and (2.15).

Let us now consider the following generalization of (2.1). There is desired in terms of $N, n, r, a_1, a_2, \cdots, a_n$, the value of

$$F_r(n, N, a_1, a_2, \cdots, a_n) = \sum \frac{N!}{x_1!\,x_2! \cdots x_n!} a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n} \tag{2.21}$$

where $a_1, a_2, \cdots, a_n$, are constants and the summation is for all values of $x_1, x_2, \cdots, x_n$ such that $x_1 + x_2 + \cdots + x_n = N$ and no $x = r$. The method of procedure is the same as that for the case already considered, viz when $a_1 = a_2 = \cdots = a_n = 1$.

The sum in (2.2) may be rearranged into the sum of a number of terms as follows:

$$(2.22)\quad \left\{\begin{array}{l}
\displaystyle\sum \frac{N!}{x_1!\,x_2! \cdots x_n!} \quad a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n}, \quad x_1 + x_2 + \cdots + x_n = N, \quad \text{no } x = r; \\
\displaystyle\frac{a_1^r}{r!} \sum \frac{N!}{x_2! \cdots x_n!} a_2^{x_2} \cdots a_n^{x_n} + \cdots + \frac{a_n^r}{r!} \sum \frac{N!}{x_1! \cdots x_{n-1}!} a_1^{x_1} \cdots a_{n-1}^{x_{n-1}}, \\
\qquad\qquad x_1 + x_2 + \cdots + x_{n-1} = N - r, \text{ etc.}, \quad \text{no } x = r; \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\displaystyle\frac{a_1^r \cdots a_k^r}{(r!)^k} \sum \frac{N!}{x_{k+1}! \cdots x_n!} a_{k+1}^{x_{k+1}} \cdots a_n^{x_n} + \cdots \\
\qquad\qquad \displaystyle + \frac{a_{n-k+1}^r \cdots a_n^r}{(r!)^k} \sum \frac{N!}{x_1! \cdots x_{n-k}!} a_1^{x_1} \cdots a_{n-k}^{x_{n-k}}, \\
\qquad\qquad x_1 + x_2 + \cdots + x_{n-k} = N - kr, \text{ etc.}, \quad \text{no } x = r; \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{array}\right.$$

For convenience, let us write

$$(2.23)\quad \left\{\begin{array}{l}
A(n, N) = (a_1 + a_2 + \cdots + a_n)^N \\
A_i(n - 1, N) = (a_1 + \cdots + a_{i-1} + a_{i+1} + \cdots + a_n)^N \\
A_{ij}(n - 2, N) = (a_1 + \cdots + a_{i-1} + a_{i+1} + \cdots + a_{j-1} + a_{j+1} + \cdots + a_n)^N \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
G_r(n, N) = F_r(n, N, a_1, a_2, \cdots, a_n) \\
G_r(n - 1, N, a_i) = F_r(n - 1, N, a_1, a_2, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n) \\
G_r(n - 2, N, a_i, a_j) = F_r(n - 2, N, a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_{j-1}, a_{j+1}, \cdots, a_n) \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{array}\right.$$

so that (2.2) may be written as

$$(2.24)\quad \begin{array}{l}
A(n, N) = G_r(n, N) + \displaystyle\frac{N^{(r)}}{r!} \sum_{i=1}^{n} a_i^r G_r(n - 1, N - r, a_i) \\
\qquad + \displaystyle\frac{N^{(2r)}}{\ } \sum^{n} a_i^r a_j^r G_r(n - 2, N - 2r, a_i, a_j) + \cdots \qquad (i \neq j, \text{ etc.})
\end{array}$$

From (2.24), there are obtained n equations

$$A_i(n-1, N-r) = G_r(n-1, N-r, a_i) + \frac{(N-r)^{(r)}}{r!} \tag{2.25}$$
$$a_j^r G_r(n-2, N-2r, a_i, a_j) + \cdots \qquad (i = 1, 2, \cdots, n, j \neq 1)$$

Multiplying (2.25) by $a_i^r N^{(r)}/r!$ and subtracting the result from (2.24), there is obtained

$$A(n, N) - \sum_{i=1}^{n} \frac{a_i^r N^{(r)}}{r!} A_i(n-1, N-r) = G_r(n, N) \tag{2.26}$$
$$- \frac{N^{(2r)}}{2!\,(r!)^2} \sum_{i,j=1}^{n} a_i^r a_j^r G_r(n-2, N-2r, a_i, a_j) - \cdots \qquad (i \neq j, \text{etc.}).$$

Continuing this procedure, there is finally obtained

$$G_r(n, N) = F_r(n, N, a_1, a_2, \cdots, a_n) = A(n, N) - \frac{N^{(r)}}{r!} \tag{2.27}$$
$$\sum_{i=1}^{n} a_i^r A_i(n-1, N-r) + \frac{N^{(2r)}}{2!\,(r!)^2} \sum_{i,j=1}^{n} a_i^r a_j^r A_{ij}(n-2, N-2r) - \cdots$$
$$(i \neq j, \text{etc.})$$

Similar results are obtainable for

$$G_{rs \cdots t} = F_{rs \cdots t}(n, N, a_1, a_2, \cdots, a_n) = \sum \frac{N!}{x_1!\, x_2! \cdots x_n!} a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n} \tag{2.28}$$

where the summation is for all values of x_i such that $x_1 + x_2 + \cdots + x_n = N$, and no $x = r, s, \cdots$, or t.

Thus, it will be shown later (see section 8), that

$$G_r(n, N) = G_{rs}(n, N) + \frac{N^{(s)}}{s!} \sum_{i=1}^{n} a_i^s G_{rs}(n-1, N-s, a_i) \tag{2.29}$$
$$+ \frac{N^{(2s)}}{2!\,(s!)^2} \sum_{i,j=1}^{n} a_i^s a_j^s G_{rs}(n-2, N-2s, a_i, a_j) + \cdots \qquad (i \neq j, \text{etc.})$$

Corresponding to the derivation of (2.27), there is obtained from (2.29) the fact that

$$G_{rs}(n, N) = G_r(n, N) - \frac{N^{(s)}}{s!} \sum_{i=1}^{n} a_i^s G_r(n-1, N-s, a_i) \tag{2.30}$$
$$+ \frac{N^{(2s)}}{2!\,(s!)^2} \sum_{i,j=1}^{n} a_i^s a_j^s G_r(n-2, N-2s, a_i, a_j) - \cdots \qquad (i \neq j, \text{etc.})$$

3. **The problem to be studied.** Consider a trial in which one of n mutually exclusive events may occur, with the respective probabilities of occurrence

$p_1, p_2, \cdots, p_n$ where $p_1 + p_2 + \cdots + p_n = 1$. The probabilities of the various combinations of events which are possible in N trials are given by the terms of the expansion of $(p_1 + p_2 + \cdots + p_n)^N$.

In the N trials some of the possible events may not occur, others may occur one, twice, etc. It is desired to study the distribution of the number of events which do not occur; the distribution of the number of events which occur once each, etc. The simultaneous distributions of the events above described are also to be studied.

For example, the possible event may be the occurrence of a digit. A study of a sequence of random digits, in sets of ten, yielded the following three sample sets.

0	1	2	3	4	5	6	7	8	9
1	0	2	1	1	2	1	0	0	2
1	1	1	1	1	1	2	0	1	1
0	0	2	1	2	1	2	1	0	1

FIG. 1

In the first set three events do not occur, four occur once each, and three occur twice each. In the second set one event does not occur, eight events occur once each, and one event occurs twice; etc.

4. **Distribution of the number of events not occurring.** To obtain the distribution of the number of events which do not occur, there is applied to the expansion of $(p_1 + p_2 + \cdots + p_n)^N$ a procedure similar to that employed in section 2.

Thus, if π_{r0} represents the probability for r events not occurring, then

$$
(4.1)\quad \left\{
\begin{aligned}
\pi_{00} &= \sum \frac{N!}{x_1!\,x_2!\cdots x_n!}\,p_1^{x_1}p_2^{x_2}\cdots p_n^{x_n}, \quad x_1 + x_2 + \cdots + x_n = N, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{no } x = 0; \\
\pi_{10} &= \sum \frac{N!}{x_2!\cdots x_n!}\,p_2^{x_2}\cdots p_n^{x_n} + \cdots + \sum \frac{N!}{x_1!\cdots x_{n-1}!}\,p_1^{x_1}\cdots p_{n-1}^{x_{n-1}}, \\
&\qquad\qquad x_1 + x_2 + \cdots + x_{n-1} = N, \text{ etc.}, \quad \text{no } x = 0; \\
\pi_{r0} &= \sum \frac{N!}{x_{r+1}!\cdots x_n!}\,p_{r+1}^{x_{r+1}}\cdots p_n^{x_n} + \cdots + \sum \frac{N!}{x_1!\cdots x_{n-r}!}\,p_1^{x_1}\cdots p_{n-r}^{x_{n-r}}, \\
&\qquad\qquad x_1 + x_2 + \cdots + x_{n-r} = N, \text{ etc.}, \quad \text{no } x = 0;
\end{aligned}
\right.
$$

Employing (2.21), we may write (4.1) as

$$(4.2)\quad \begin{cases} \pi_{00} = F_0(n, N, p_1, p_2, \cdots, p_n) \\ \pi_{10} = F_0(n-1, N, p_2, \cdots, p_n) + \cdots + F_0(n-1, N, p_1, p_2, \cdots, p_{n-1}) \\ \pi_{r0} = F_0(n-r, N, p_{r+1}, \cdots, p_n) + \cdots + F_0(n-r, N, p_1, \cdots, p_{n-r}) \end{cases}$$

Since $p_1 + p_2 + \cdots + p_n = 1$ there is found from (2.27) that

$$(4.3)\quad \begin{aligned} \pi_{00} &= 1 - \sum_{i=1}^{n} (1-p_i)^N + \frac{1}{2!} \sum_{i,j=1}^{n} (1-p_i-p_j)^N \\ &\qquad - \frac{1}{3!} \sum_{i,j,k=1}^{n} (1-p_i-p_j-p_k)^N + \cdots \\ \pi_{10} &= \sum_{i=1}^{n} (1-p_i)^N - \sum_{i,j=1}^{n} (1-p_i-p_j)^N \\ &\qquad + \frac{1}{2!} \sum_{i,j,k=1}^{n} (1-p_i-p_j-p_k)^N - \cdots \\ \pi_{20} &= \frac{1}{2!}\left\{ \sum_{i,j=1}^{n} (1-p_i-p_j)^N - \sum_{i,j,k=1}^{n} (1-p_i-p_j-p_k)^N + \cdots \right\} \\ \pi_{30} &= \frac{1}{3!}\left\{ \sum_{i,j,k=1}^{n} (1-p_i-p_j-p_k)^N - \cdots \right\} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (i \neq j, \text{ etc.}) \end{aligned}$$

The factorial moments[5] of the distribution given by (4.3) are easily derived. The first factorial moment is given by $\sigma_1 = \pi_{10} + 2\pi_{20} + 3\pi_{30} + \cdots + r\pi_{r0} + \cdots$ and the summation of the proper terms in (4.3) yields

$$(4.4)\quad \sigma_1 = \sum_{i=1}^{n} (1-p_i)^N$$

In general, the r-th factorial moment, given by $\sigma_r = \sum_{k=r}^{n} k(k-1) \cdots (k-r+1)\pi_{k0}$ is

$$(4.5)\quad \sigma_r = \sum_{a,b,\cdots,r=1}^{n} (1-p_a-p_b-\cdots-p_r)^N, \qquad (a \neq b, \text{ etc.}).$$

Indeed, (4.3) illustrates the fact that, if $f(x)$ is the probability that a discontinuous variate takes the value x, then[6]

$$(4.6)\quad f(x) = \frac{1}{x!} \sum_{k=0}^{n-x} (-1)^k \sigma_{x+k}/k!$$

[5] J. F. Steffensen, *Interpolation* (1927), p. 101.

[6] J. F. Steffensen, "Factorial Moments and Discontinuous Frequency Functions" *Skandinavisk Aktuarietidskrift*, Vol. VI (1923), pp. 73–89.

The moments about any constant of the distribution given by (4.3) may be derived from the factorial moments by the relation[7]

$$(4.7)\quad E(x-a)^r = (1 + \sigma_1\Delta + \sigma_2\Delta^2/2! + \cdots + \sigma_r\Delta^r/r!)\cdot\xi^r \qquad (\xi = -a)$$

where Δ is the difference operator of the calculus of finite differences, and ξ is replaced by $(-a)$ after the indicated operations have been performed.

Of special interest is the case when $p_1 = p_2 = \cdots = p_n = \dfrac{1}{n}$, for which (4.3) becomes

$$(4.8)\quad \begin{aligned} &= \left(\frac{1}{n}\right)^N f_0(n, N) = \left(\frac{1}{n}\right)^N \Delta^n 0^N \\ &= \left(\frac{1}{n}\right)^N n f_0(n-1, N) = \left(\frac{1}{n}\right)^N n\Delta^{n-1}0^N \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \pi_{r0} &= \left(\frac{1}{n}\right)^N \binom{n}{r} f_0(n-r, N) = \left(\frac{1}{n}\right)^N \binom{n}{r}\Delta^{n-r}0^N \end{aligned}$$

where $f_0(n, N)$ and $\Delta^n 0^N$ are as defined in section 2. The probabilities in (4.8) are the respective terms of the expansion of $\left(\dfrac{1}{n}\right)^N (1+\Delta)^n\cdot 0^N$.

For this case the r-th factorial moment becomes

$$(4.9)\quad \sigma_r = n(n-1)\cdots(n-r+1)\,(n-r)^N/n^N$$

There is presented an example of the distribution (4.8) for the case $n = N = 10$. It is found that[8]

$$(4.10)\quad \begin{cases} \Delta 0^{10} = 1 & \Delta^6 0^{10} = 16435440 \\ \Delta^2 0^{10} = 1022 & \Delta^7 0^{10} = 29635200 \\ \Delta^3 0^{10} = 55980 & \Delta^8 0^{10} = 30240000 \\ \Delta^4 0^{10} = 818520 & \Delta^9 0^{10} = 16329600 \\ \Delta^5 0^{10} = 5103000 & \Delta^{10} 0^{10} = 3628800 \end{cases}$$

$$(4.11)\quad \begin{array}{llll} \pi_{00} = .000362880 & & \pi_{50} = .128595600 \\ \pi_{10} = .016329600 & & \pi_{60} = .017188920 \\ \pi_{20} = .136080000 & & \pi_{70} = .000671760 \\ \pi_{30} = .355622400 & & \pi_{80} = .000004599 \\ \pi_{40} = .345144240 & & \pi_{90} = .000000001 \end{array}$$

$$(4.12)\quad \begin{array}{llll} \sigma_1 = 3.486784401 & & m = 3.486784401 \\ \sigma_2 = 9.663676416 & & \sigma^2 = 0.992795358 \end{array}$$

[7] This result is derived as follows: $(x-a)^r = (1+\Delta)^x\cdot(-a)^r$; $E(x-a)^r = \sum_{x=1}^{n}(x-a)^r f(x) = \left(\sum_{x=1}^{n}(1+\Delta)^x\cdot f(x)\right)\cdot(-a)^r = \left(\sum_{x=1}^{n}(1 + x\Delta + x(x-1)\Delta^2/2! + \cdots)f(x)\right)\cdot(-a)^r$. For a bivariate distribution it may be shown similarly that, symbolically, $E((x-a)^r(y-b)^s) = \{\exp(\sigma_{1.}\,\Delta_1 + \sigma_{.1}\,\Delta_2)\}\cdot(-a)^r(-b)^s$ where $\sigma_{1.}{}^m\sigma_{.1}{}^n = \sigma_{mn}$ and Δ_1 operates only on a and Δ_2 operates only on b. A similar result may be derived for a multivariate distribution.

[8] cf. Whittaker & Robinson, *op. cit.* p. 7.

The observed distribution was obtained by distributing 200 sets of ten digits each, the digits being found in Tippet's Random Sampling Numbers.[9] The results obtained are given in Fig. 2. Three of the 200 observed sets were illustrated in section 3.

The agreement between observed results and theoretical values is gratifying.

5. **Distribution of the number of events which occur once each.** Let π_{k1}, represent the probability that there are k events which occur once each. Thus, the various probabilities, obtained by rearranging the terms of the expansion of $(p_1 + p_2 + \cdots + p_n)^N$, are as follows:

$$\pi_{01} = \sum \frac{N!}{x_1! \cdots x_n!} p_1^{x_1} \cdots p_n^{x_n}, \quad x_1 + x_2 + \cdots + x_n = N, \quad \text{no } x = 1;$$

(5.1)

$$\pi_{11} = p_1 \sum \frac{N!}{x_2! \cdots x_n!} p_2^{x_2} \cdots p_n^{x_n} + \cdots + p_n \sum \frac{N!}{x_1! \cdots x_{n-1}!} p_1^{x_1} \cdots p_{n-1}^{x_{n-1}},$$
$$x_1 + x_2 + \cdots + x_{n-1} = N - 1, \text{ etc.}, \quad \text{no } x = 1;$$

$$\pi_{k1} = p_1 p_2 \cdots p_k \sum \frac{N!}{x_{k+1}! \cdots x_n!} p_{k+1}^{x_{k+1}} \cdots p_n^{x_n} + \cdots + p_{n-k+1} \cdots p_n$$
$$\sum \frac{N!}{x_1! \cdots x_{n-k}!} p_1^{x_1} \cdots p_{n-k}^{x_{n-k}},$$
$$x_1 + x_2 + \cdots + x_{n-k} = N - k, \text{ etc.}, \quad \text{no } x = 1;$$

No. of events not occurring x	Observed frequency f	Theoretical frequency	xf	$x(x-1)f$	Observed parameters
0	0	0.08	0	0	$\bar{\sigma}_1 = 3.46$
1	8	3.26	8	0	$\bar{\sigma}_2 = 9.61$
2	22	27.22	44	44	$\bar{x} = 3.46$
3	72	71.12	216	432	$s^2 = 1.0984$
4	72	69.02	288	864	Theoretical
5	21	25.72	105	420	Parameters
6	4	3.44	24	120	$\sigma_1 = 3.49$
7	1	0.14	7	42	$\sigma_2 = 9.66$
8	0	0.00	0	0	$m = 3.49$
9	0	0.00	0	0	$\sigma^2 = 0.99$
	200	200.00	692	1922	

FIG. 2

[9] L. H. C. Tippet, Random Sampling Numbers, *Tracts for Computers*, No. XV (1927), London.

In view of (2.21) and (2.27), it is found that (5.1) becomes

$$
(5.2)\quad
\begin{aligned}
\pi_{01} &= 1 - N\sum_{i=1}^{n} p_i(1-p_i)^{N-1} + \frac{N(N-1)}{2!}\sum_{i,j=1}^{n} p_ip_j(1-p_i-p_j)^{N-2} - \cdots \\
\pi_{11} &= N\left\{\sum_{i=1}^{n} p_i(1-p_i)^{N-1} - (N-1)\sum_{i,j=1}^{n} p_ip_j(1-p_i-p_j)^{N-2} + \cdots\right\} \\
\pi_{21} &= \frac{N(N-1)}{2!}\left\{\sum_{i,j=1}^{n} p_ip_j(1-p_i-p_j)^{N-2} - \cdots\right\} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (i \neq j, \text{ etc.})
\end{aligned}
$$

From (5.2) there is readily derived the fact that

$$
(5.3)\quad
\begin{aligned}
\sigma_r = N(N-1)\cdots(N-r+1) \\
\sum_{a,b,\cdots,r=1}^{n} p_ap_b\cdots p_r(1-p_a-p_b-\cdots-p_r)^{N-r}, \qquad (a \neq b, \text{ etc.})
\end{aligned}
$$

For the case in which $p_1 = p_2 = \cdots = p_n = \frac{1}{n}$, the distribution in (5.2) becomes

$$
(5.4)\quad
\begin{cases}
\pi_{01} = \left(\frac{1}{n}\right)^N f_1(n, N) \\
\phantom{\pi_{01}} = \left(\frac{1}{n}\right)^N nNf_1(n-1, N-1) \\
\pi_{21} = \left(\frac{1}{n}\right)^N \frac{n(n-1)N(N-1)}{2!} f_1(n-2, N-2) \\
\pi_{r1} = \left(\frac{1}{n}\right)^N \binom{n}{r} N^{(r)} f_1(n-r, N-r) \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{cases}
$$

where $f_1(n, N)$ and $N^{(r)}$ have been defined in section 2. For this case (5.3) becomes

$$
(5.5)\qquad \sigma_r = n^{(r)}N^{(r)}(n-r)^{N-r}/n^N
$$

Evaluation of (5.4) and (5.5) for $n = N = 10$ yields,

$$
(5.6)\quad
\begin{cases}
\pi_{01} = .00811639 & \pi_{41} = .27052704 & \pi_{81} = .01632960 \\
\pi_{11} = .04794633 & \pi_{51} = .15621984 & \pi_{91} = .00000000^{10} \\
\pi_{21} = .14082336 & \pi_{61} = .12700800 & \pi_{101} = .00036288 \\
\pi_{31} = .21089376 & \pi_{71} = .02177280 &
\end{cases}
$$

$$
(5.7)\quad
\begin{aligned}
\sigma_1 &= 3.87420489 & m &= 3.87420489 \\
\sigma_2 &= 13.58954496 & \sigma^2 &= 2.45428632
\end{aligned}
$$

[10] For the case $n = N = 10$ there cannot be 9 events occurring once each, since then the tenth event must also occur once.

The observed distribution, given in Fig. 3, was obtained from the 200 sets previously considered.

The agreement between the observed results and theoretical values is gratifying.

6. Distribution of the number of events which occur r times each. Let π_{kr} represent the probability that there are k events occurring r times each. Thus, the various probabilities, obtained by rearranging the terms of the expansion of $(p_1 + p_2 + \cdots + p_n)^N$, are as follows:

No. of events occurring once each x	Observed frequency f	Theoretical frequency	xf	$x(x-1)f$	Observed parameters
0	1	1.62	0	0	$\bar{\sigma}_1 = 3.905$
1	10	9.58	10	0	$\bar{\sigma}_2 = 14.000$
2	30	28.16	60	60	$\bar{x} = 3.905$
3	37	42.18	111	222	$s^2 = 2.656$
4	62	54.10	248	744	Theoretical
5	27	31.24	135	540	Parameters
6	22	25.40	132	660	$\sigma_1 = 3.874$
7	3	4.36	21	126	$\sigma_2 = 13.590$
8	8	3.26	64	448	$m = 3.874$
9	0	0.00	0	0	$\sigma^2 = 2.454$
10	0	0.08	0	0	
	200	199.98	781	2800	

FIG. 3

$$\pi_{0r} = \sum \frac{N!}{x_1! \cdots x_n!} p_1^{x_1} \cdots p_n^{x_n}, \quad x_1 + x_2 + \cdots + x_n = N, \quad \text{no } x = r;$$

$$\pi_{1r} = \frac{p_1^r}{r!} \sum \frac{N!}{x_2! \cdots x_n!} p_2^{x_2} \cdots p_n^{x_n} + \cdots + \frac{p_n^r}{r!} \sum \frac{N!}{x_1! \cdots x_{n-1}!} p_1^{x_1} \cdots p_{n-1}^{x_{n-1}},$$

$$x_1 + x_2 + \cdots + x_{n-1} = N - r, \text{ etc.}, \quad \text{no } x = r;$$

(6.1) ....

$$\pi_{kr} = \frac{p_1^r p_2^r \cdots p_k^r}{(r!)^k} \sum \frac{N!}{x_{k+1}! \cdots x_n!} p_{k+1}^{x_{k+1}} \cdots p_n^{x_n} + \cdots$$

$$+ \frac{p_{n-k+1}^r \cdots p_n^r}{(r!)^k} \sum \frac{N!}{x_1! \cdots x_{n-k}!} p_1^{x_1} \cdots p_{n-k}^{x_{n-k}},$$

$$x_1 + x_2 + \cdots + x_{n-k} = N - kr, \text{ etc.}, \quad \text{no } x = r;$$

In view of (2.21) and (2.27) it is found that (6.1) becomes

$$
(6.2)\quad
\begin{aligned}
\pi_{0r} &= 1 - \frac{N^{(r)}}{r!}\sum_{i=1}^{n} p_i^r(1-p_i)^{N-r} + \frac{N^{(2r)}}{2!\,(r!)^2}\sum_{i,j=1}^{n} p_i^r p_j^r(1-p_i-p_j)^{N-2r} - \cdots \\
\pi_{1r} &= \frac{N^{(r)}}{r!}\left\{\sum_{i=1}^{n} p_i^r(1-p_i)^{N-r} - \frac{(N-r)^{(r)}}{r!}\sum_{i,j=1}^{n} p_i^r p_j^r(1-p_i-p_j)^{N-2r} + \cdots\right\} \\
\pi_{2r} &= \frac{N^{(2r)}}{2!\,(r!)^2}\left\{\sum_{i,j=1}^{n} p_i^r p_j^r(1-p_i-p_j)^{N-2r} - \cdots\right\} \\
&\cdots\cdots\cdots\cdots\cdots\cdots \qquad (i \neq j, \text{ etc.})
\end{aligned}
$$

From (6.2) there is readily derived the fact that

$$
(6.3)\quad \sigma_k = \frac{N^{(kr)}}{(r!)^k}\sum_{a,b,\cdots,k=1}^{n} p_a^r p_b^r \cdots p_k^r(1-p_a-p_b-\cdots-p_k)^{N-kr}, \quad (a \neq b, \text{etc.})
$$

For $r = 0,1$ (6.2) and (6.3) reduce to the values previously derived.

For the case in which $p_1 = p_2 = \cdots = p_n = \dfrac{1}{n}$, the distribution in (6.2) becomes

$$
(6.4)\quad
\begin{aligned}
\pi_{0r} &= \left(\frac{1}{n}\right)^N f_r(n, N) \\
\pi_{1r} &= \left(\frac{1}{n}\right)^N \frac{nN^{(r)}}{r!} f_r(n-1, N-r) \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
{}_{:r} &= \left(\frac{1}{n}\right)^N \binom{n}{k} \frac{N^{(kr)}}{(r!)^k} f_r(n-k, N-kr)
\end{aligned}
$$

where $f_r(n, N)$ has been defined in section 2. For this case (6.3) becomes

$$
(6.5)\quad \sigma_k = N^{(kr)} n^{(k)} (n-k)^{N-kr}/n^N
$$

7. Simultaneous distribution of the number of events not occurring, and of the number of events occurring once each. The probabilities for the simultaneous occurrence of the various combinations of the number of events not occurring, and of the number of events occurring once each, are given by rearranging the terms of the expansion of $(p_1 + p_2 + \cdots + p_n)^N$, and are given as in Fig. 4.

In Fig. 4 none of the subscripts take on equal values simultaneously, and G_{01} has been defined in section 2. Summation of the values in the k-th column of Fig. 4, yields the probability that there are $(k-1)$ events not occurring. Comparison with (4.2) yields

$$
(7.1)\quad
\begin{aligned}
F_0(n, N, p_1, p_2, \cdots, p_n) &= G_0(n, N) = G_{01}(n, N) + N\sum_{i=1}^{n} p_i G_{01}(n-1, N-1, p_i) \\
&+ \frac{N^{(2)}}{2!}\sum_{i,j=1}^{n} p_i p_j G_{01}(n-2, N-2, p_i, p_j) + \cdots, \quad (i \neq j, \text{ etc.})
\end{aligned}
$$

		Number of events not occurring		
		0	1	r
Number of events occurring once each	0	$G_{01}(n, N)$	$\sum_{i=1}^{n} G_{01}(n-1, N, p_i)$	$\cdots$
	1	$N\sum_{i=1}^{n} p_i G_{01}(n-1, N-1, p_i)$	$N\sum_{i,j=1}^{n} p_i G_{01}(n-2, N-1, p_i, p_j)$	$\cdots$
	2	$\frac{N^{(2)}}{2!}\sum_{i,j=1}^{n} p_i p_j G_{01}(n-2, N-2, p_i p_j)$	$\frac{N^{(2)}}{2!}\sum_{i,j,k=1}^{n} p_i p_j G_{01}(n-3, N-2, p_i, p_j, p_k)$	$\cdots$
	s	$\cdots$	$\cdots$	$\frac{N^{(s)}}{r!\,s!}\sum_{a,b,\cdots,s,\alpha,\beta,\cdots,\rho=1}^{n} p_a p_b \cdots p_s G_{01}(n-r-s, N-s, p_a, \cdots, p_s, p_\alpha, \cdots, p_\rho)$

FIG. 4

Summation of the values in the k-th row of Fig. 4, yields the probability that there are $(k-1)$ events occurring once each. Comparison with (5.2) and (2.27) yields

$$
\begin{aligned}
F_1(n, N, p_1, p_2, \cdots, p_n) &= G_1(n, N) = G_{01}(n, N) + \sum_{i=1}^{n} G_{01}(n-1, N, p_i) \\
&\quad + \frac{1}{2!}\sum_{i,j=1}^{n} G_{01}(n-2, N, p_i, p_j) + \cdots, \quad (i \neq j, \text{ etc.})
\end{aligned} \tag{7.2}
$$

If we use x to represent the number of events not occurring, and y the number of events occurring once each, then it is found that

$$
\begin{aligned}
E(x^{(r)}y^{(s)}) = \sigma_{rs} = N^{(s)} &\sum_{a,b,\cdots,s,\alpha,\beta,\cdots,\rho=1}^{n} p_a p_b \cdots p_s(1 - p_a - \cdots - p_s \\
&- p_\alpha - \cdots - p_\rho)^{N-s}, \quad (a \neq b, \text{ etc.}).
\end{aligned} \tag{7.3}
$$

If ${}_0\bar{x}_{k1}$ represents the average number of events not occurring, when there are k events occurring once each, then from Fig. 4 there is found that

$$
{}_0\bar{x}_{01} = \frac{\displaystyle \sum_{i=1}^{n} G_{01}(n-1, N, p_i) + 2\sum_{i,j=1}^{n} G_{01}(n-2, N, p_i, p_j)/2! + 3\sum_{i,j,k=1}^{n} G_{01}(n-3, N, p_i, p_j, p_k)/3! + \cdots}{\displaystyle G_{01}(n, N) + \sum_{i=1}^{n} G_{01}(n-1, N, p_i) + \sum_{i,j=1}^{n} G_{01}(n-2, N, p_i, p_j)/2! + \cdots} \quad (i \neq j, \text{ etc.}) \tag{7.4}
$$

In view of (7.2), (7.4) reduces to

$$(7.5) \qquad {}_0\bar{x}_{01} = \left(\sum_{i=1}^{n} G_1(n, N, p_i)\right) \Big/ G_1(n, N)$$

A similar procedure, yields, in general

$$(7.6) \quad {}_0\bar{x}_{kl} = \frac{\displaystyle\sum_{a,b,\cdots,k,l=1}^{n} p_a p_b \cdots p_k G_1(n-k-1, N-k, p_a, p_b, \cdots, p_k, p_l)}{\displaystyle\sum_{a,b,\cdots,k=1}^{n} p_a p_b \cdots p_k G_1(n-k, N-k, p_a, p_b, \cdots, p_k)} \qquad (a \neq b, \text{ etc.})$$

If ${}_1\bar{y}_{k0}$ represents the average number of events occurring once each, when there are k events not occurring, then from Fig. 4, there is found that

$$(7.7) \quad {}_1\bar{y}_{00} = \frac{N\left\{\displaystyle\sum_{i=1}^{n} p_i G_{01}(n-1, N-1, p_i) + 2(N-1) \sum_{i,j=1}^{n} p_i p_j G_{01}(n-2, N-2, p_i, p_j)/2! + \cdots\right\}}{G_{01}(n, N) + N \displaystyle\sum_{i=1}^{n} p_i G_{01}(n-1, N-1, p_i) + N^{(2)} \sum_{i,j=1}^{n} p_i p_j G_{01}(n-2, N-2, p_i, p_j)/2!} \qquad (i \neq j, \text{ etc.})$$

In view of (7.1), (7.7) reduces to

$$(7.8) \qquad {}_1\bar{y}_{00} = \left(N \sum_{i=1}^{n} p_i G_0(n-1, N-1, p_i)\right) \Big/ G_0(n, N)$$

A similar procedure, yields, in general

$$(7.9) \quad {}_1\bar{y}_{k0} = \frac{N \displaystyle\sum_{a,b,\cdots,k,l=1}^{n} p_a G_0(n-k-1, N-1, p_a, p_b, \cdots, p_k, p_l)}{\displaystyle\sum_{a,b,\cdots,k=1}^{n} G_0(n-k, N, p_a, p_b, \cdots, p_k)} \qquad (a \neq b, \text{ etc.})$$

For the case in which $p_1 = p_2 = \cdots = p_n = \dfrac{1}{n}$, as may be found from Fig. 4, the probability for the simultaneous occurrence of r events not occurring, and s events occurring once each, is given by

$$(7.10) \qquad \left(\frac{1}{n}\right)^N \frac{n^{(r+s)} N^{(s)}}{r!\, s!} f_{01}(n-r-s, N-s)$$

For this case (7.1), (7.2), (7.3), (7.6), and (7.9) yield respectively

$$(7.11) \quad f_0(n, N) = f_{01}(n, N) + nN f_{01}(n-1, N-1) + \binom{n}{2} N^{(2)} f_{01}(n-2, N-2) + \cdots$$

$$(7.12) \quad f_1(n, N) = f_{01}(n, N) + n f_{01}(n-1, N) + \binom{n}{2} f_{01}(n-2, N) + \cdots$$

$$(7.13) \quad \sigma_{rs} = N^{(s)} n^{(r+s)} (n-r-s)^{N-s}/n^N$$

(7.14) $\quad {}_0\bar{x}_{k1} = (n - k)f_1(n - k - 1, N - k)/f_1(n - k, N - k)$

(7.15) $\quad {}_1\bar{y}_{k0} = N(n - k)f_0(n - k - 1, N - 1)/f_0(n - k, N)$

Let us consider again the case when $p_1 = p_2 = \cdots = p_n = \dfrac{1}{n}$ and $n = N = 10$. Evaluating (7.14) and (7.15) by means of (2.15) yields

$$\left\{\begin{array}{llll} {}_0\bar{x}_{01} = 5.71 & & {}_0\bar{x}_{51} = 3.02 \\ {}_0\bar{x}_{11} = 5.21 & & {}_0\bar{x}_{61} = 2.10 \\ {}_0\bar{x}_{21} = 4.51 & & {}_0\bar{x}_{71} = 2.00 \\ {}_0\bar{x}_{31} = 4.10 & & {}_0\bar{x}_{81} = 1.00 \\ {}_0\bar{x}_{41} = 3.28 & & {}_0\bar{x}_{91} = 0.00 \end{array}\right. \tag{7.16}$$

$$\left\{\begin{array}{llll} {}_1\bar{y}_{00} = 10.00 & & {}_1\bar{y}_{50} = 1.83 \\ {}_1\bar{y}_{10} = 8.00 & & {}_1\bar{y}_{60} = 0.89 \\ {}_1\bar{y}_{20} = 6.16 & & {}_1\bar{y}_{70} = 0.27 \\ {}_1\bar{y}_{30} = 4.50 & & {}_1\bar{y}_{80} = 0.02 \\ {}_1\bar{y}_{00} = 3.05 & & {}_1\bar{y}_{90} = 0.00 \end{array}\right. \tag{7.17}$$

The 200 sets of observations already considered yielded the simultaneous distribution given in Fig. 5.

Number of events occurring once each	Number of events not occurring: 0	1	2	3	4	5	6	7	8	9		$\bar{x}$
0								1			1	7.00
1					1	6	3				10	5.20
2					16	13	1				30	4.50
3					35	2					37	4.05
4				42	20						62	3.32
5				27							27	3.00
6			19	3							22	2.14
7			3								3	2.00
8		8									8	1.00
9											0	
10											0	
	0	8	22	72	72	21	4	1	0	0	200	
$\bar{y}$		8.00	6.16	4.46	3.03	1.81	1.25	0.00				

FIG. 5

The distribution in Fig. 5 yields $\bar{\sigma}_{11} = 11.89$, (7.13) yields $\sigma_{11} = 12.07959552$.

The agreement between the observed results in Fig. 5 and the theoretical values in (7.16) and (7.17) is gratifying.

8. Simultaneous distribution of the number of events which occur r times each, and of the number of events which occur s times each. The probabilities for the simultaneous occurrence of the various combinations of the number of events which occur r times each, and of the number of events which occur s times each, are obtained by rearranging the terms of the expansion of $(p_1 + p_2 + \cdots + p_n)^N$. If $\pi_{kr,ls}$ is the probability for the simultaneous occurrence of k events which occur r times each and l events which occur s times each, then

$$\pi_{kr,ls} = \frac{N^{(kr+ls)}}{k!\, l!\, (r!)^k (s!)^l} \sum_{a,b,\cdots,k,\alpha,\beta,\cdots,\lambda=1}^{n} p_a^r \cdots p_k^r p_\alpha^s \cdots p_\lambda^s G_{rs} (n - k - l, N - kr - ls, p_a, \cdots, p_k, p_\alpha, \cdots, p_\lambda), \qquad (a \neq b, \text{etc.}) \tag{8.1}$$

where G_{rs} is defined in section 2.

From (8.1) and (6.2), there is derived, in a manner similar to the derivation of (7.1) and (7.2), the result that

$$F_r(n, N, p_1, \cdots, p_n) = G_r(n, N) = G_{rs}(n, N) + \frac{N^{(s)}}{s!} \sum_{i=1}^{n} p_i^s G_{rs}(n-1, N-s, p_i) + \frac{N^{(2s)}}{2!\,(s!)^2} \sum_{i,j=1}^{n} p_i^s p_j^s G_{rs}(n-2, N-2s, p_i, p_j) + \cdots, \qquad (i \neq j, \text{etc.}) \tag{8.2}$$

and a similar result by interchanging r and s in (8.2).

For the distribution given by (8.1), it is found that

$$\sigma_{kl} = \frac{N^{(kr+ls)}}{(r!)^k (s!)^l} \sum_{a,b,\cdots,k,\alpha,\beta,\cdots,\lambda=1}^{n} p_a^r \cdots p_k^r p_\alpha^s \cdots p_\lambda^s (1 - p_a - \cdots - p_k - p_\alpha - \cdots - p_\lambda)^{N-kr-ls}, \qquad (a \neq b, \text{etc.}) \tag{8.3}$$

If ${}_r\bar{x}_{ls}$ represents the average number of events which occur r times each when there are l events which occur s times each, then from (8.1) and (8.2), in a manner similar to the derivation of (7.6), it is found that

$${}_r\bar{x}_{ls} = \frac{(N - ls)^{(r)} \sum_{a,\alpha,\cdots,\lambda=1}^{n} p_a^r p_\alpha^s \cdots p_\lambda^s G_s(n - 1 - l, N - r - ls, p_a, p_\alpha, \cdots, p_\lambda)}{r! \sum_{\alpha,\cdots,\lambda=1}^{n} p_\alpha^s \cdots p_\lambda^s G_s(n - l, N - ls, p_\alpha, \cdots, p_\lambda)} \qquad (\alpha \neq \beta, \text{etc.}) \tag{8.4}$$

If ${}_s\bar{y}_{kr}$ represents the average number of events which occur s times each when there are k events which occur r times each, then by interchanging k and l, and r and s in (8.4), there is found

$$
{}_s\bar{y}_{kr} = \frac{(N-kr)^{(s)} \sum\limits_{a,\cdots,k,\alpha=1}^{n} p_a^r \cdots p_k^r p_\alpha^s G_r(n-k-1, N-kr-s, p_a, \cdots, p_k, p_\alpha)}{\sum\limits_{a,b,\cdots,k=1}^{n} p_a^r \cdots p_k^r G_r(n-k, N-kr, p_a, \cdots, p_k)} \tag{8.5}
$$

$(a \neq b, \text{etc.})$

For the case when $p_1 = p_2 = \cdots = p_n = \dfrac{1}{n}$, it is found that (8.1), (8.2), (8.3), (8.4), and (8.5) respectively yield

$$
\pi_{kr,ls} = \left(\frac{1}{n}\right)^N \frac{n^{(k+l)} N^{(kr+ls)}}{k!\, l!\, (r!)^k\, (s!)^l} f_{rs}(n-k-l, N-kr-ls) \tag{8.6}
$$

$$
\begin{aligned}
f_r(n, N) = f_{rs}(n, N) &+ \frac{nN^{(s)}}{s!} f_{rs}(n-1, N-s) \\
&+ \frac{n(n-1)N^{(2s)}}{2!\,(s!)^2} f_{rs}(n-2, N-2s) + \cdots
\end{aligned} \tag{8.7}
$$

$$
\sigma_{kl} = n^{(k+l)} N^{(kr+ls)} (n-k-l)^{N-kr-ls} / (r!)^k\, (s!)^l\, n^N \tag{8.8}
$$

$$
{}_r\bar{x}_{ls} = (n-l)(N-ls)^{(r)} f_s(n-1-l, N-r-ls)/r!\, f_s(n-l, N-ls) \tag{8.9}
$$

$$
{}_s\bar{y}_{kr} = (n-k)(N-kr)^{(s)} f_r(n-k-1, N-kr-s)/s!\, f_r(n-k, N-kr) \tag{8.10}
$$

For $r = 0$, $s = 1$, the results derived in this section of course reduce to those already derived in section 7.

9. **Conclusion.** It is clear that the same method of procedure may be employed to study the simultaneous distribution of the number of events which occur r, s, $\cdots$, t, times each. However we will not continue the discussion any further.

We have thus seen that the multinomial distribution serves as the background for the study of a number of distributions which have certain practical applications.

The theory discussed herein has been illustrated by several examples which yielded gratifying agreement between observed and theoretical results.

Washington, D. C.

A PROBLEM IN LEAST SQUARES

By Jan K. Wiśniewski

§1. We are dealing with two variables, the observed values of which are denoted x and y respectively. The pairs of observations are divided into r groups, numbering $n_1, n_2, \cdots n_r$ pairs. Suppose in each group we determine a regression equation of the following shape:

$$y_i = a_i + b_i x + \cdots m_i x^s \tag{1}$$

where y_i denotes the value of the "dependent" variable obtained from the regression equation, while y without any subscript denotes its observed value. The r regression equations of type (1) are not assumed independent; on the contrary, we postulate that

$$\sum_1^r y_i = a_0 + b_0 x + \cdots m_0 x^s \tag{2}$$

be fulfilled identically in x; $a_0, b_0, \cdots m_0$ being predetermined numbers. This leads to the following conditions:

$$\sum_1^r a_i = a_0 \qquad \sum_1^r b_i = b_0 \cdots \sum_1^r m_i = m_0. \tag{3}$$

The magnitude to be minimized under the theory of least squares is now

$$Z = \sum_1^{r-1} \sum\nolimits_i [y - (a_i + b_i x + \cdots m_i x^s)]^2 + \sum\nolimits_r \left\{ y - \left[\left(a_0 - \sum_1^{r-1} a_i \right) + \left(b_0 - \sum_1^{r-1} b_i \right) x + \cdots \left(m_0 - \sum_1^{r-1} m_i \right) x^s \right] \right\}^2. \tag{4}$$

The normal equations derived from (4) are of the following shape:

$$n_j a_j + n_r \sum_1^{r-1} a_i + b_j \sum\nolimits_j x + \left(\sum_1^{r-1} b_i \right) \left(\sum\nolimits_r x \right) + \cdots m_j \sum\nolimits_j x^s + \left(\sum_1^{r-1} m_i \right) \left(\sum\nolimits_r x^s \right) = \sum\nolimits_j y - \sum\nolimits_r y + n_r a_0 + b_0 \sum\nolimits_r x + \cdots m_0 \sum\nolimits_r x^s \tag{5}$$

$$a_j \sum_j x + \left(\sum_1^{r-1} a_i\right)\left(\sum_r x\right) + b_j \sum_j x^2 + \left(\sum_1^{r-1} b_i\right)\left(\sum_r x^2\right)$$
$$+ \cdots m_j \sum_j x^{s+1} + \left(\sum_1^{r-1} m_i\right)\left(\sum_r x^{s+1}\right) = \sum_j xy - \sum_r xy + a_0 \sum_r$$
$$+ b_0 \sum_r x^2 + \cdots m_0 \sum_r x^{s+1}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (5)$$

$$a_j \sum_j x^s + \left(\sum_1^{r} a_i\right)\left(\sum_r x^s\right) + b_j \sum_j x^{s+1} + \left(\sum_1^{r-1} b_i\right)\left(\sum_r x^{s+1}\right)$$
$$+ \cdots m_j \sum_j x^{2s} + \left(\sum_1^{r-1} m_i\right)\left(\sum_r x^{2s}\right) = \sum_j x^s y - \sum_r x^s y$$
$$+ a_0 \sum_r x^s + b_0 \sum_r x^{s+1} + \cdots m_0 \sum_r x^{2s}$$

$\sum_i$ meaning a summation extended over the i-th group. As (1) is of the s-th degree, we have $(s + 1)$ $(r - 1)$ parameters to determine and as many equations, the problem thus being in theory solved.* As to the numerical solution, Doolittle's method or any other may be applied. We do not enter at present the question, how much labor would the actual solution require.

Examples. Allen and Bowley in their book on "Family Expenditure" (London, 1935) assume the expenditure on some defined item f to be a linear function of the total expenditure e

$$f = ke + c. \qquad (6)$$

Evidently $\sum k = 1$, $\sum c = 0$ (cfr. pp. 10–11). Another example I give in a paper on seasonal variation, which appeared in "Economic Studies" III (Kraków). Actual values y of a time series are assumed to be linear functions of certain "normal" values x

$$y = a + bx \qquad (7)$$

a and b changing from month to month but constant from year to year. Then $\sum a = 0$, $\sum b = 12$.

§2. Methods of solution in special cases. The generally recognized methods of solving normal equations become extremely laborious as the product $(s + 1)$ $(r - 1)$ grows large. As a matter of fact, the amount of computer's work is approximately proportional to the cube of the number of parameters to determine. Therefore short cuts seem to be indispensable. A most elegant one is at our disposal in the special case[1] when the values of x in the several groups

* The remaining $s + 1$ parameters a_r, b_r, $\cdots$ m are, of course, found from (3).

[1] This seems to be realized in Allen and Bowley's work.

are identical, or, at least, the sums n_i, $\sum_i x$, $\sum_i x^2$, $\cdots$ $\sum_i x^{2s}$ are identical in i. Instead of (1) we shall write

$$y_i = A_i + B_i X_1 + \cdots M X_s \tag{8}$$

where $X_1, X_2, \cdots X_s$ are orthogonal polynomials, i.e. such that $\sum X_i X_j = 0$ if and only if $i \neq j$. In general, $X_h = X^h + \alpha_{h-1}^h X^{h-1} + \cdots \alpha_0^h$, the coefficients being rational functions of n, $\sum x$, $\sum x^2$, $\cdots$ $\sum x^{2s-1}$.

The conditions (3) can now be replaced by a set of equivalent ones, viz.

$$\sum_1^r A_i = A_0 \qquad \sum_1^r B_i = B_0 \cdots \sum_1^r M_i = M_0. \tag{9}$$

How the actual values of A_0, B_0, $\cdots$ M_0 are found, will be shown in the next paragraph. The solution becomes now very easy, as the normal equations for the determination of each set of $r - 1$ parameters are independent, i.e. we can calculate the A's separately, then the B's etc., the order of solution being of no importance. Moreover the shape of the normal equations permits of considerable simplification of solution. Suppose we have to determine the values of the coefficients K, corresponding to X_h. The normal equations are now—after certain simplifications—

$$\begin{aligned}
2K_1 + K_2 + K_3 + \cdots K_{r-1} &= \frac{1}{\sum X_h^2}\left(\textstyle\sum_1 X_h y - \sum_r X_h y\right) + K_0 \\
K_1 + 2K_2 + K_3 + \cdots K_{r-1} &= \frac{1}{\sum X_h^2}\left(\textstyle\sum_2 X_h y - \sum_r X_h y\right) + K_0 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
K_1 + K_2 + K_3 + \cdots 2K_{r-1} &= \frac{1}{\sum X_h^2}\left(\textstyle\sum_{r-1} X_h y - \sum_r X_h y\right) + K_0.
\end{aligned} \tag{10}$$

Adding these equations, dividing the sum by r and substracting the quotient from the j-th equation, we get

$$K_j = \frac{\sum_j X_h y}{\sum X_h^2} - \frac{1}{r}\left(\sum_1^r \frac{\sum_i X_h y}{\sum X_h^2} - K_0\right). \tag{11}$$

The first member of the right hand side of (11) should be regarded as the principal term: this is actually the value we would obtain for K_j, were this coefficient independent from the other K's. The second member is a correction term, the necessary amount of correction being distributed equally among the several K's. The simple solution given by (11) is only possible if the sum $\sum X_h^2$ is the same for each group. From the definition of X_h we see that it is equivalent to saying that n_i, $\sum_i x$, $\sum_i x^2$, $\cdots$ $\sum_i x^{2h}$ be identical in i. As h increases to s, we come to the condition given at the beginning of this paragraph.

§3. If this condition is not fulfilled, we can, indeed, replace the power series in x by orthogonal polynomials $X_{h \cdot i}$, the second subscript being appended in order to show that the values of the X polynomials are no more identical for the several groups; these polynomials are now orthogonalized separately within each group. But we are no more able to predetermine the values of $A_0, B_0, \cdots M_0$, as they depend on each other; this will be made clear a little later. Therefore we have to resort to an approximation: the values of the parameters will not be found from simultaneous equations, but successively, step by step, beginning with those corresponding to the highest degree of the independent variable.

The values of $a_0, b_0, \cdots m_0$ are given. It is evident that $m_0 = M_0$. The j-th normal equation is now:

$$M_j \sum_j X_{s \cdot j}^2 - M_0 \sum_r X_{s \cdot r}^2 + \left(\sum_1^{r-1} M_i \right) \left(\sum_r X_{s \cdot r}^2 \right) = \sum_j X_{s \cdot j} y - \sum_r X_{s \cdot j} y. \quad (12)$$

We see at once that

$$M_i = \frac{M_j \sum_j X_{s \cdot j}^2 + \sum_i X_{s \cdot i} y - \sum_j X_{s \cdot j} y}{\sum_i X_{s \cdot i}^2}. \quad (13)$$

Inserting this into /12/ we get

$$M_j = \frac{\sum_j X_{s \cdot j} y}{\sum_j X_{s \cdot j}^2} - \frac{1}{\sum_j X_{s \cdot j}^2} \cdot \frac{\sum_1^r \frac{\sum_i X_{s \cdot i} y}{\sum_i X_{s \cdot i}^2} - M_0}{\sum_1^r \sum_i X_{s \cdot i}^2}. \quad (14)$$

The second member of the right hand side of /14/ is again a correction term, the necessary amount of correction being distributed in inverse proportion to $\sum_j X_{s \cdot j}^2$. Now we determine the value of L_0, this coefficient corresponding to $s - 1$, the second highest degree of x, and calculate the several L's from equations strictly analogous to (14) thus accomplishing the second step of our work, and so on, down to the A's. L_0 is found from the following equation:

$$L_0 = l_0 - \sum_1^r [\alpha_{s-1}^s(i) \cdot M_i]. \quad (15)$$

To α_{s-1}^s is now appended a bracketed i, this to stress its variation from group to group. We see from (15) that before the several M's are calculated we are not in a position to determine L_0. On the other hand, if α_{s-1}^s is the same for all groups, the second member of the right hand side of (15) simply reduces to $\alpha_{s-1}^s \cdot m_0$ and L_0 can be determined in advance, i.e. before calculating the M's. This is the case treated first (in §2). In any case, if no definite correlation is to be expected between $\alpha_{s-1}^s(i)$ and M_i, the approximative method developed here should give very nearly correct results. The writer applied this method of solution to the simple problem of seasonal variation mentioned in §1 and found the results very satisfactory.

A SIGNIFICANCE TEST FOR COMPONENT ANALYSIS

By Paul G. Hoel

1. Introduction

During the last few years several papers and books have been written on various aspects of what has been termed component or factor analysis. This analysis has arisen from the psychological problem of describing the results on a series of tests in terms of a few distinct abilities or components. In much of such work it is claimed that there does not exist more than a certain number of components, the material discarded in order to substantiate such a claim being considered as due to random errors of sampling or errors of measurement. However, mere inspection of results or the calculation of standard errors of residual correlations is hardly sufficient to justify such conclusions, and therefore a significance test of some kind is necessary. Hotelling[1] considered such a test but based it upon an uncertain analogy with the analysis of variance and upon the legitimacy of using standard errors. The purpose of this paper is to derive a test which is more general in scope and in which all assumptions are explicitly stated.

If each test score is thought of as being made up of two parts, a true score and an error element, the assumption that there exists fewer components than the number of tests implies that the scatter diagram of the true scores will lie in a space of correspondingly smaller dimensionality. Consequently, an ideal test for the number of components would be one which would test the rank of the true moment matrix. In the case of normally distributed variables, this line of approach leads one to the sampling distribution of the generalized variance. Unfortunately, this distribution appears in unintegrated form; however, by considering its moments it is possible to find a good approximation to this exact distribution for samples which are not too small.

The paper proceeds by first finding two approximation distributions for the generalized variance, one for samples which are not too small and one for large samples. It then considers the type of population from which it will be assumed the sample was drawn, and finally applies the test to two numerical examples from recent literature along such lines.

2. Approximation Distributions

Suppose that N individuals have been drawn at random from an n variate normal population whose distribution is expressed by

$$(1) \qquad P(x_1, x_2, \cdots, x_n) = Ke^{-\sum_1^n A_{ij} x_i x_j}$$

[1] Harold Hotelling, Analysis of a Complex of Statistical Variables into Principal Components, The Journal of Educational Psychology, September and October, 1933, pp. 21–25.

where $x_i = X_i - m_i$, $A_{ij} = \dfrac{\Delta_{ij}}{2\sigma_i\sigma_j\Delta}$, Δ is the determinant $|\rho_{ij}|$ and Δ_{ij} is the cofactor of ρ_{ij} in Δ, and $K = |A_{ij}|^{\frac{1}{2}}/(2\pi)^{n/2}$. If the observed values of the variables of the αth individual are denoted by $X_{i\alpha}(i = 1, 2, \cdots, n)$, then the generalized sample variance is defined as $z = |a_{ij}|$, where $a_{ij} = \dfrac{1}{N}\sum_{\alpha=1}^{N}(X_{i\alpha} - \bar{X}_i)(X_{j\alpha} - \bar{X}_j)$. Wilks[2] has shown that in sampling from the population (1), the kth moment of the sampling distribution of z is given by

$$M_k \quad 1^{-k} \frac{\Gamma\left(\frac{N+2k-1}{2}\right)\Gamma\left(\frac{N+2k-2}{2}\right)\cdots\Gamma\left(\frac{N+2k-n}{2}\right)}{\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\cdots\Gamma\left(\frac{N-n}{2}\right)}$$

where $A = N^n|A_{ij}|$. An inspection of the integrated form of the distribution of z in the case of $n = 1$ and $n = 2$ suggests that there likely exists a function of similar form for higher values of n whose kth moment can be made to differ from M_k only in higher powers of terms which contain N^{-1} as a factor. An investigation along such lines leads to the function

$$g(z) = Cz^m e^{-n\sqrt[n]{az}} \tag{2}$$

where $C = \dfrac{a^{\frac{N-n}{2}} n^{n\frac{N-n}{2}-1}}{\Gamma\left(n\frac{N-n}{2}\right)}$, $m = \dfrac{N-n-2}{2}$, $a = Aq$ and $q = 1 - \dfrac{(n-1)(n-2)}{2N}$.

It will be shown that the kth moment M_k' of $g(z)$ differs from M_k only in terms of magnitude less than the second and higher powers of k^2n/N or kn^2/N.

Multiplying $g(z)$ by z^k and integrating over the entire range of z will yield M_k', which turns out to be

$$M_k' = \frac{\Gamma\left(n\frac{N-n+2k}{2}\right)}{a^k n^{nk}\Gamma\left(n\frac{N-n}{2}\right)}.$$

Upon reducing the upper gamma function and performing successive steps of simple algebra

$$\begin{aligned} M_k' &= a^{-k}n^{-nk}\left(n\frac{N-n+2k}{2} - 1\right)\left(n\frac{N-n+2k}{2} - 2\right)\cdots\left(n\frac{N-n}{2}\right) \\ &= N^{nk}a^{-k}2^{-nk}\left(1 + \frac{2k-n-2/n}{N}\right)\left(1 + \frac{2k-n-4/n}{N}\right)\cdots \\ &\qquad \left(1 + \frac{2k-n-2kn/n}{N}\right). \end{aligned}$$

[2] S. S. Wilks, Certain Generalizations in the Analysis of Variance, Biometrika, Vol. XXIV, 1923, p. 477.

The terms in parentheses may be treated as the factored form of a polynomial of the nkth degree in unity. Thus the quantities $\dfrac{2k - n - 2/n}{N}$, etc., may be treated as the zeros with signs changed of the corresponding polynomial in x (say). As a result, the successive terms after the first in the non-factored form of this polynomial in unity are the sums of the products of these quantities taken one at a time, two at a time, etc. Upon performing this multiplication and letting $\phi = N^n/2^nA$, M_k' assumes the form

$$M_k' = \phi^k q^{-k}\left[1 - \frac{k(n^2 - nk + 1)}{N} + \cdots\right]$$

where the neglected terms are in magnitude less than the second and higher powers of k^2n/N or kn^2/N. If M_k is handled in exactly the same manner, it will be found that

$$\begin{aligned}M_k &= A^{-k}\left(\frac{N + 2k - 1}{2} - 1\right)\cdots\left(\frac{N + 2k - 1}{2} - k\right)\cdots \\ &\qquad\qquad \left(\frac{N + 2k - n}{2} - 1\right)\cdots\left(\frac{N + 2k - n}{2} - k\right) \\ &= N^{nk}A^{-k}2^{-nk}\left(1 + \frac{2k - 3}{N}\right)\cdots\left(1 - \frac{1}{N}\right)\cdots \\ &\qquad\qquad \left(1 + \frac{2k - n - 2}{N}\right)\cdots\left(1 - \frac{n}{N}\right) \\ &= \phi^k\left[1 - \frac{nk(n - 2k + 3)}{2N} + \cdots\right]\end{aligned}$$

where the neglected terms are of the same order of magnitude as those neglected in the approximation to M_k'. Before a comparison of M_k and M_k' is possible, the factor q^{-k} of M_k' must be expanded and multiplied into the quantity in brackets. This operation yields the result

$$M_k' = \phi^k\left[1 - \frac{nk(n - 2k + 3)}{2N} + \cdots\right].$$

Thus M_k and M_k' agree to within neglected terms. As a matter of fact, if the values of the neglected terms are considered more carefully, it will be found that the actual difference between M_k and M_k' is considerably less than the given upper bound for the magnitude of neglected terms would indicate. For example, when $n = 5$ the first term in the difference is $6k(k - .9)N^{-2}$, while $625k^2N^{-2}$ or $25k^4N^{-2}$ is the upper bound for this term when only general results are used. The general formula for the first term in this difference has been obtained, but since the remaining terms have not been investigated and since the type of problems to which the distribution $g(z)$ is to be applied does not

justify this refinement, it will not be considered here. Consequently, if one considers this distribution function as sufficiently determined by its low order moments and if one applies $g(z)$ only to problems in which N is fairly large compared with n^2, then the function $g(z)$ will give a good approximation to the exact sampling distribution of z. Obviously, $g(z)$ is identical with the exact distribution for the known cases of $n = 1$ and $n = 2$. It is not possible under the above expansions to vary the constants in the form of $g(z)$ in such a manner as to obtain an approximation whose kth moment will agree with M_k to within still higher powers of comparable terms.

In order to test whether or not a sample value $z = Z$ can be reasonably assumed to have been obtained in random sampling from a population of type (1) with fixed A, it is necessary to calculate the probability P of obtaining in repeated samples a value of z greater than Z. Thus it is necessary to evalua te

$$1 - \int_0^Z g(z)\,dz.$$

Upon making the substitution $x = n\sqrt[n]{az}$, and letting $p = n\dfrac{N-n}{2} - 1$ and $u = n\sqrt[n]{aZ}\left(n\dfrac{N-n}{2}\right)^{-\frac{1}{2}} = nN\sqrt[n]{\dfrac{Z}{\phi}\left[1 - \dfrac{(n-1)(n-2)}{2N}\right]}\,[2n(N-n)]^{-\frac{1}{2}}$, this integral can be reduced to the standard form of the incomplete gamma function. Hence P assumes the form

$$P = 1 - I(u, p) \tag{3}$$

where

$$I(u, P) = \frac{1}{\Gamma(p+1)} \int_0^{u\sqrt{p+1}} e^{-x} x^p\,dx.$$

In many applications of this distribution it will be found that the values of u and p lie beyond the tabled[8] values of these constants. Consequently, it will often be sufficient to use the normal distribution to which the gamma distribution tends as N becomes large. This normal distribution will be considered next.

Rather than obtain a normal approximation to $g(z)$ or the gamma function to which $g(z)$ reduces after the above transformation, it is more illuminating to find the basic descriptive parameters of the exact distribution of z and from them obtain a normal approximation. Such a procedure will show how rapidly the distribution of z approaches normality with increasing N. By using the recurrence formula connecting M_{k+1} and M_k, which can be found directly from the ratio of these two moments, and expressing the necessary moments in

[8] K. Pearson, Tables of the Incomplete Gamma Function, Biometric Laboratory (1922), Univ. of London.

terms of M_1, it can be shown that these basic descriptive parameters are expressible in expanded form as follows:

$$m = \phi\left[1 - \frac{n(n+1)}{2N} + \frac{n(n+1)(n-1)(3n+2)}{24\,N^2} + \cdots\right]$$

$$\sigma^2 = \phi^2\left[\frac{2n}{N} - \frac{n(2n^2 - n + 1)}{N^2} + \cdots\right]$$

$$\beta_1 = \frac{2(3n-1)^2}{nN}\left[1 - \frac{(n+1)(5n-3)}{2(3n-1)N} + \cdots\right]$$

$$\beta_2 = 3\left[1 + \frac{4(3n-1)(4n-1)}{3nN} + \cdots\right].$$

These values suggest that

$$w = \sqrt{\frac{N}{2n}}\left[\frac{z}{\phi} - 1\right] \tag{4}$$

will likely be distributed approximately normally with zero mean and unit variance. As a matter of fact, by using the second limit theorem of probability,[4] it can be shown that the distribution of w approaches normality as N increases indefinitely. Hence, for samples in which N is large compared with n^2, it will be sufficient to compare the value of w arising from a sample $z = Z$ with its variance of unity if a test of significance is desired. A better general approximation could have been obtained by centering the curve at $\phi\left[1 - \frac{n(n+1)}{2N}\right]$ rather than at ϕ; however, since there is positive skewness and the true mean lies between these two values, there might arise some exaggeration in a significance test in doing so because the accuracy of such a test depends upon the accuracy of the approximation in the right hand tail of the curve.

Inspection of (3) and (4) shows that the only population parameter upon which these approximation distributions depend is ϕ. There are no assumptions necessary about the population means, or variances, or covariances, except in so far as they may be related when the value of ϕ is postulated. This means that either (3) or (4) enables one to test whether or not it is reasonable to assume that the sample variance $z = Z$ arose in random sampling from some normal population with ϕ equal to the postulated value.

3. Population Assumptions

Consider the set of variables $u_1, u_2, \cdots, u_n$ distributed according to the normal law

$$P(u_1, u_2, \cdots, u_n) = K_1 e^{-\sum_1^n b_{ij} u_i u_j} \tag{5}$$

[4] See, for example, Frechet and Shohat, A Proof of the Generalized Second Limit Theorem in the Theory of Probability, Transactions of the American Mathematical Society, Vol. 33, (1931), p. 533.

and the set of variables $v_1, v_2, \cdots, v_n$ distributed according to the normal law

$$(6) \qquad P(v_1, v_2, \cdots, v_n) = K_2 e^{-\sum_1^n c_i v_i^2}$$

where the v's are uncorrelated with the u's and with each other. The joint distribution of the u's and v's is expressed by

$$(7) \qquad P(u_1, \cdots, v_n) = K_3 e^{-\sum_1^n b_{ij} u_i u_j - \sum_1^n c_i v_i^2}.$$

Upon writing down the determinant of the coefficients of these $2n$ variables, it will become evident that any one of its principal minors of any order can be expressed as the product of a principal minor of $| b_{ij} |$ with a principal minor of $| c_i |$. Since the distributions (5) and (6) are normal, the determinants $| b_{ij} |$ and $| c_i |$ are positive definite; consequently the determinant of the coefficients in (7) must also be positive definite.

Now consider the orthogonal transformation

$$y_i = \frac{u_i + v_i}{\sqrt{2}}, \qquad i = 1, 2, \cdots, n$$

$$y_i = \frac{u_i - v_i}{\sqrt{2}}, \qquad i = n + 1, \cdots, 2n.$$

Since the determinant of the coefficients in (7) is invariant under an orthogonal transformation, the resulting distribution of the y's may be expressed by

$$(8) \qquad P(y_1, y_2, \cdots, y_{2n}) = K_4 e^{-\sum_1^{2n} d_{ij} y_i y_j}$$

where $| d_{ij} |$ is positive definite.

In order to obtain the distribution of the variables $y_1, y_2, \cdots, y_n$, it is necessary to integrate (8) with respect to the variables $y_{n+1}, \cdots, y_{2n}$ over their range of values. If this integration is performed after the quadratic form in the exponent of (8) has been expressed as a sum of squares[5] with coefficients which are the ratios of principal minors of $| d_{ij} |$, it will be clear that the integration leaves a quadratic form in the exponent which is also positive definite. Hence after the transformation $x_i = \sqrt{2} y_i (i = 1, 2, \cdots, n)$ the distribution function of the variables $x_i = u_i + v_i (i = 1, 2, \cdots, n)$ must be normal and may be expressed by (1). Thus it has been shown that if the true parts u_i of the variables x_i are normally distributed without error and if the error parts v_i are normally distributed but are uncorrelated with the u_i and with each other, then the variables x_i possess a normal distribution. The advantage of

[5] See, for example, Risser and Traynard, Les Principes de la Statistique Mathematique, 1933, p. 225.

this formulation will become evident when the parameter ϕ is expressed in terms of the parameters of (5) and (6).

Since the v's are uncorrelated with the u's and with each other, the variance σ_i^2 of x_i is the sum of the variances of u_i and v_i, while the correlation ρ_{ij} between x_i and x_j may be expressed in terms of the correlation ρ'_{ij} between u_i and u_j and the variances u_i^2, u_j^2, ν_i^2, ν_j^2 of u_i, u_j, v_i, v_j respectively. These relationships are

$$(9) \qquad \sigma_i^2 = \mu_i^2 + \nu_i^2, \quad \text{and} \quad \rho_{ij} = \frac{\rho'_{ij}}{\sqrt{(1 + \nu_i^2/\mu_i^2)(1 + \nu_j^2/\mu_j^2)}} \qquad (i \neq j).$$

For simplicity of notation let $\lambda_i = \nu_i^2/\mu_i^2$. Now it is well known[6] that ϕ can be expressed in the form

$$\phi = \sigma_1^2 \sigma_2^2 \cdots \sigma_n^2 \mid \rho_{ij} \mid .$$

If the values from (9) are inserted in $\mid \rho_{ij} \mid$ and if the resulting denominators of elements are factored out, ϕ will assume the form

$$\phi = \frac{\sigma_1^2 \sigma_2^2 \cdots \sigma_n^2 B}{(1 + \lambda_1) \cdots (1 + \lambda_n)}$$

where

$$B = \begin{vmatrix} 1 + \lambda_1 & \rho'_{12} & \cdots & \rho'_{1n} \\ \rho'_{12} & & & \\ & & & \\ \rho'_{1n} & \cdots & \cdots & 1 + \lambda_n \end{vmatrix}$$

Following the methods of confluence analysis,[7] B can be expressed as follows:

$$B = R + \sum_{\alpha=1}^{n} \lambda_\alpha R_{)\alpha(} + \sum_{\alpha<\beta}^{n} \lambda_\alpha \lambda_\beta R_{)\alpha\beta(} + \cdots + \lambda_1 \lambda_2 \cdots \lambda_n$$

where $R = \mid \rho'_{ij} \mid$, $R_{)\alpha(}$ is the principal minor of R obtained by deleting row and column α, etc. R is the true correlation determinant whose rank it is the object of this paper to test. If R is assumed to be of rank $n - t$, then all principal minors containing more than $n - t$ rows vanish and B reduces to

$$B = \sum_{\alpha_1 < \cdots < \alpha_t}^{n} \lambda_{\alpha_1} \lambda_{\alpha_2} \cdots \lambda_{\alpha_t} R_{)\alpha_1 \alpha_2 \cdots \alpha_t(} + \cdots + \lambda_1 \lambda_2 \cdots \lambda_n .$$

The tests (3) and (4) were designed to test hypothetical values of ϕ by means of the sample Z. Evidently the value of ϕ can be postulated by assigning hypothetical values to the λ's, the σ's, and the principal minors of R.

Assigning values to the λ's does not curtail the degrees of freedom in these

---

[6] S. S. Wilks, loc. cit., p. 477.

[7] Ragnar Frisch, Statistical Confluence Analysis by Means of Complete Regression Systems, Oslo, 1934.

tests because they were derived on the basis of (1) which depends only on the m's, σ's, and ρ's. The λ's do restrict the range of the ρ's, but not their degrees of freedom.

An inspection of the expression for ϕ shows that ϕ can be made to assume any desired value irregardless of the rank of R by merely assigning the σ's properly. It is therefore necessary to make some assumption regarding the σ's if the test is to serve the purpose for which it is intended. Here it will be sufficient to assume that the product of the population variances may be replaced by the product of the sample variances. This assumption will ordinarily be approximately fulfilled for the size samples for which it is legitimate to employ (3) or (4); consequently this assumption does not restrict the range of application of the test.

To postulate values of the principal minors of R beyond postulating the rank of R would introduce hypotheses and restrictions which are irrelevant to the fundamental purpose of the test. This difficulty will be avoided by replacing all non-vanishing minors of R by their upper bounds of unity. Since this will overestimate the value of B, and hence of ϕ, the usual significance level of .05 may be considered as decisive. Let the value of B when unity is inserted for all non-vanishing principal minors be denoted by D. Then

$$D = \sum_{\alpha_1 < \cdots < \alpha_t}^{n} \lambda_{\alpha_1} \lambda_{\alpha_2} \cdots \lambda_{\alpha_t} + \cdots + \lambda_1 \lambda_2 \cdots \lambda_n . \tag{10}$$

Since

$$\prod_1^n (1 + \lambda_i) = 1 + \sum_{\alpha=1}^n \lambda_\alpha + \sum_{\alpha_1 < \alpha_2}^n \lambda_{\alpha_1} \lambda_{\alpha_2} + \cdots + \lambda_1 \lambda_2 \cdots \lambda_n$$

it will often be convenient to write D in the form

$$D = \prod_1^n (1 + \lambda_i) - \left\{ 1 + \sum_{\alpha=1}^n \lambda_\alpha + \cdots + \sum_{\alpha_1 < \cdots < \alpha_{t-1}}^n \lambda_{\alpha_1} \lambda_{\alpha_2} \cdots \lambda_{\alpha_{t-1}} \right\}. \tag{11}$$

As a consequence of all the above assumptions,

$$\begin{aligned} \frac{Z}{\phi} = \frac{| a_{ij} |}{\phi} &= \frac{(1 + \lambda_1) \cdots (1 + \lambda_n) \, | r_{ij} |}{B} \\ &\geqq \frac{(1 + \lambda_1) \cdots (1 + \lambda_n) \, | r_{ij} |}{D} \end{aligned} \tag{12}$$

where $| r_{ij} |$ is the sample correlation determinant.

All the essential material for testing the rank of the true correlation matrix is contained in (3), (4), (11), and (12). In summary, the hypothesis to be tested and the procedure to follow in performing the test are as follows.

The population of n variables from which the sample is supposed drawn is assumed to be such that (a) the true parts of the variables are normally distributed, (b) the error parts are normally distributed but are uncorrelated with the true parts and with each other, (c) the product of the variances may be replaced by the product of the sample variances, (d) the values of the λ's

are postulated as judged by the accuracy in measurement of the variables, and (e) the rank of the true correlation matrix is $n - t$.

Given the value $|r_{ij}|$ of the sample correlation determinant, a lower bound for the value of Z/ϕ is calculated from (11) and (12). This lower bound is inserted in either (3) or (4), depending on the size of the sample. If (3) is used and if $P \leqq .05$, or if (4) is used and $w \geqq 2$, one may conclude, as judged by the sample variance, that it is very unlikely that the sample was drawn in random sampling from the population specified above. If one has reason to believe that the variables are sensibly normal as indicated above and that the postulated values of the λ's are quite accurate, then the test shows quite definitely that the postulated rank of the true correlation matrix is unsubstantiated by the sample, and therefore a higher rank should be tested until a non-significant value is obtained. Because a lower bound rather than the value of Z/ϕ is used, the test can be used on minimum ranks only, and hence a value of $Z < \phi$ will not yield a test of significance. However, the test does handle the problem for which it was designed and which is of fundamental interest, and that is to see whether or not one is justified in assuming that a sample represents only a certain minimum number of components.

4. Applications

(a) Hotelling[8] has used an example taken from other sources to illustrate his test on components. In order to compare results, this same example will be treated here under the assumptions outlined above. In this example the reliability coefficients are given. From the definition of a reliability coefficient r_i, it follows at once that $r_i = \dfrac{1}{1 + \lambda_i}$. The population values of the λ's will be set equal to the values obtained from these sample reliability coefficients. The data for this problem are

$$|r_{ij}| = .235,\ N = 140,\ n = 4,\ \lambda_1 = .087,\ \lambda_2 = .119,\ \lambda_3 = .101,\ \lambda_4 = .773.$$

Assume that the true correlation matrix in the population is of rank two, that is, that two components are sufficient to describe the results on these tests. Since N is large compared with n^2, it will be sufficient to use (4). The values of (11), (12), and (4) are found to be

$$D = \prod^{4} (1 + \lambda_i) - \left\{1 + \sum^{4} \lambda_\alpha\right\} = .294$$

$$\frac{Z}{\phi} \geqq \frac{\prod (1 + \lambda_i) \, |r_{ij}|}{D} = 1.90$$

$$w \geqq \sqrt{\frac{140}{8}}\,[1.90 - 1] = 3.76$$

[8] Loc. cit., p. 16.

Since the standard deviation of w is unity, this value demonstrates clearly that the hypothesis of only two components is untenable as judged by the sample correlation determinant. If one assumes three components, the test will be found to yield a non-significant value. Hence it may be concluded that under the hypotheses on which the test is based, the sample does not justify the assumption of less than three components. Hotelling's test indicated the necessity for two components but was uncertain about the third, the decision resting upon a variate value of 1.31 as against a standard deviation of unity.

(b) Thurstone, in his "Vectors of Mind," considers an example taken from a series of fifteen psychological tests. After applying his centroid method to the data, he inspects his results and concludes that four components are sufficient to account for everything except random errors. It is impossible to test his conclusions explicitly as above because the size of the sample is not given and the reliability coefficients are not known. Nevertheless, if it is legitimate to assume that the sample is sufficiently large to justify the use of this test, interesting conclusions can be obtained on the assumption that only four components are needed.

Suppose that $\lambda_i = \frac{1}{2}$, which implies that the variance of error is half as large as the true sampling variance for each variable. Here (10) is more convenient than (11) for computing the value of D. The values of (10) and (12) are found to be

$$D = {}_{15}C_3(\tfrac{1}{2})^{12} + {}_{15}C_2(\tfrac{1}{2})^{13} + {}_{15}C_1(\tfrac{1}{2})^{14} + (\tfrac{1}{2})^{15} = .125$$

$$\frac{Z}{\phi} \geq \frac{|r_{ij}|}{.0003}.$$

Evidently, the value of $|r_{ij}|$ must lie in the neighborhood of .0003 if the test is not to yield a significant result which contradicts the hypothesis. However, the correlations in $|r_{ij}|$ are given to only three decimal places, and therefore a legitimate value in the neighborhood of .0003 can not be realized. It is to be noted that the postulated values of the λ's are equivalent to postulating that all reliability coefficients are equal to $\frac{2}{3}$, a value which should be considered as unusually low. It would seem reasonable to avoid using material in which the variance of error is larger than one-half the variance of random sampling, unless the variance of random sampling is exceedingly small.

CONTRIBUTIONS TO THE THEORY OF COMPARATIVE STATISTICAL ANALYSIS. I. FUNDAMENTAL THEOREMS OF COMPARATIVE ANALYSIS[1]

By William G. Madow

This is the first of several papers in which there will be presented a general approach to the statistical examination of hypotheses which are false if any of several things are true. Phenomena requiring such a statistical theory are investigated quite frequently. As examples may be cited the studies of lag correlation in time series, periodogram analysis in geophysics, factor analysis in psychology, and analysis into components in agriculture.[2]

The theorems of this paper have one purpose: to permit the reduction of the distributions by which the hypotheses are to be tested to essentially the joint distribution of the statistics which contain the information offered by the data concerning the truth or falsity of the things which will negate the hypotheses. In order to do this it has been necessary to generalize the theorem of Poincare on the probability that at least one of several events occur.[3] As illustrations there are stated, after Theorems III, VI, and IX, generalizations of a distribution derived by Jordan, (5) page 109.[4]

In a second paper, we shall give a complete derivation of the joint distributions necessary for the applications of the analysis of variance. A reconsideration of the Schuster periodogram will be included. In other papers these results will be extended to problems arising in the theory of regression, and to problems of the distributions of medians, etc.

The fundamental theorems of comparative analysis are now obtained in such a form that they are applicable to problems in the theory of probability no matter what the distributions may be. Some special cases of these theorems[5]

[1] Presented to the American Mathematical Society, March 27, 1937. Research under a grant-in-aid from the Carnegie Corporation of New York.

[2] Naturally these techniques are also useful in other branches of science then those in which they were first applied. It should be noted that by analysis into components we here refer to the work of Fisher, (2), chapter 6.

[3] See, Poincaré, (7), page 60. This theorem is attributed to Poincaré by Jordan, (5), and Fréchet, (3).

[4] This distribution states the probability that in r trials of an experiment which has exactly n possible results, these results being mutually exclusive, each of the possible results occurs at least once. Jordan's derivation has been simplified by Fréchet, (3), page 12.

[5] The theorems are, of course, part of the theory of measure and integration.

have been used in connection with the derivation of distributions of positional statistics such as the k^{th} in order of N elements,[6] and others.

Let Ω be a collection of elements x, and let Δ be a set of subsets of Ω. Then, the axioms which the elements of Δ are to satisfy are[7]

I. Δ is a field;[8]

II. $\Omega \epsilon \Delta$;

III. To every $A \epsilon \Delta$ there is ordered a non-negative real number $P(A)$;

IV. $P(\Omega) = 1$;

V. If $A \epsilon \Delta$ and $B \epsilon \Delta$, and $AB = 0$, then $P(A + B) = P(A) + P(B)$.

We shall regard Ω as the set of possible results of an experiment ϵ. By events we shall mean elements of Δ. The complement $\bar{A}$ of A with respect to Ω will be an element of Δ if A is an element of Δ. $\bar{A}$ consists of all elements of Ω which are not elements of A and hence is the event which occurs if and only if A does not occur.[9]

Let the subsets of Ω

$$E_1, E_2, \cdots, E_k \tag{1}$$

be elements of Δ. Then, if $\alpha_1, \alpha_2, \cdots, \alpha_k$ is a permutation of $1, 2, \cdots, k$, the set

$$E_{\alpha_1} E_{\alpha_2} \cdots E_{\alpha_j} \bar{E}_{\alpha_{j+1}} \cdots \bar{E}_{\alpha_k} \tag{2}$$

is an element of Δ and is the event which occurs whenever all the events $E_{\alpha_1}, E_{\alpha_2}, \cdots, E_{\alpha_j}$ occur, while none of the events $E_{\alpha_{j+1}}, E_{\alpha_{j+2}}, \cdots, E_{\alpha_k}$ occur.

The events (1) are said to be independent if and only if

$$P(E_{\alpha_1} \cdots E_{\alpha_j} \bar{E}_{\alpha_{j+1}} \cdots \bar{E}_{\alpha_k}) = \prod_{\nu=1}^{j} P(E_{\alpha_\nu}) \cdot \prod_{\nu=j+1}^{k} P(\bar{E}_{\alpha_\nu}) \tag{3}$$

for all selections of the sets (1) and their complements.[10]

Theorem I. The probability that the first j of the k events (1) *occur, while the remaining $k - j$ events do not occur, is*

[6] See, for example, Gumbel, (4). It is noted that Theorems I, II, and III are stated by Arne Fisher, (1), page 42, who assumes, however, that the events are independent.

[7] These axioms are stated by Kolmogoroff, (6), page 2.

[8] A set of sets is a field if the fact that A and B are elements of the set implies that $A + B$, AB, and $A - AB$ are also elements of the set.

[9] The event A will be said to have occurred if the result of the performance of the experiment E is an element of A.

[10] See Kolmogoroff, (6), page 9 for a discussion of various equivalent definitions of independence.

$$(4)\quad P(E_1 \cdots E_j\bar{E}_{j+1} \cdots \bar{E}_k) = \sum_{\nu=0}^{k-j} (-1)^\nu \sum_{\substack{\alpha_1,\cdots,\alpha_\nu=j+1\\ \alpha_1<\alpha_2<\cdots<\alpha_\nu}}^{k} P(E_1 \cdots E_jE_{\alpha_1} \cdots E_{\alpha_\nu}).$$[11]

Proof. Let $k = j + 1$. Then it follows from Axiom V that

$$(5)\qquad P(E_1E_2 \cdots E_j) = P(E_1E_2 \cdots E_jE_{j+1}) + P(E_1E_2 \cdots E_j\bar{E}_{j+1}).$$

Hence the theorem is true for $k = j + 1$ and any $j > 0$. Let the theorem be true for $k = j, j + 1, \cdots, k - 1$. From Axiom V it follows that

$$(6)\quad P(E_1 \cdots E_j\bar{E}_{j+1} \cdots \bar{E}_k) = P(E_1 \cdots E_j\bar{E}_{j+1} \cdots \bar{E}_{k-1}) - P(E_1 \cdots E_j\bar{E}_{j+1} \cdots \bar{E}_{k-1}E_k).$$

Substituting from (4) the theorem is proved.

Let $n \geq n_1 + \cdots + n_t$, $n_i \geq 0$ $(i = 1, \cdots, t)$; and let

$$\frac{n!}{n_1!\, n_2! \cdots n_t!\, (n - n_1 - \cdots - n_t)!} \equiv (n; n_1, n_2, \cdots, n_t).$$

COROLLARY. If, for each value of ν, $(\nu = 1, 2, \cdots, k - j)$, the $(k - j; \nu)$ terms

$$P(E_1 \cdots E_jE_{\alpha_1} \cdots E_{\alpha_\nu})$$

which can be obtained by selecting $\alpha_1, \alpha_2, \cdots, a_\nu$ without repetition from $j + 1, j + 2, \cdots, k$, are all equal, then

$$(7)\qquad P(E_1 \cdots E_j\bar{E}_{j+1} \cdots \bar{E}_k) = \sum^{k-j} (-1)^\nu(k - j; \nu)P(E_1 \cdots E_{j+\nu}).$$

Let

$$(8)\qquad S(\nu) = \sum_{\substack{\alpha_1,\cdots,\alpha_\nu=1\\ \alpha_1<\cdots<\alpha_\nu}}^{n} P(E_{\alpha_1}E_{\alpha_2} \cdots E_{\alpha_\nu})$$

where the summation extends over the $(k; \nu)$ terms

$$(9)\qquad P(E_{\alpha_1}E_{\alpha_2} \cdots E_{\alpha_\nu})$$

which can be obtained by selecting ν of the k events (1) without repetition. If all the terms (9) which can be obtained by selecting ν of the k events (1) without repetition are equal, then

$$(10)\qquad S(\nu) = (k; \nu)P(E_1 \cdots E_\nu).$$

[11] By definition

$$\sum_{\nu=0}^{k-j} (-1)^\nu \sum_{\substack{\alpha_1,\cdots,\alpha_\nu=j+1\\ \alpha_1<\cdots<\alpha}}^{k} P(E_1 \cdots E_j E_{j+1} \cdots \bar{E}_{\alpha_\nu})$$

$$P(E_1 \cdots E_j) + \sum_{\nu=1}^{k-j} (-1)^\nu \sum_{\substack{\alpha_1,\cdots,\alpha_\nu=j+1\\ \alpha_1<\cdots<\alpha_\nu}}^{k} P(E_1 \cdots E_jE_{\alpha_1} \cdots E_{\alpha_\nu}).$$

Theorem II. The probability that exactly j of the k events (1) *occur is*

$$P_{(j)} = \sum^{k-j} (-1)^{\nu}(j + \nu; \nu)S(j + \nu). \tag{11}$$

Proof. If $A_{(j)}$ is the subset of Ω defined by the requirement that exactly j of the events (1) occur, then $A_{(j)}$ is the sum of $(k; j)$ disjunct sets:

$$A_{(j)} = \sum_{\alpha_1, \cdots, \alpha_j = 1}^{k} E_{\alpha_1} \cdots E_{\alpha_j} \bar{E}_{\alpha_{j+1}} \cdots \bar{E}_{\alpha_k}, \tag{12}$$

where $\alpha_{j+1}, \cdots, \alpha_k$ have those of the values $1, \cdots, k$ which remain after the selection of $\alpha_1, \cdots, \alpha_j$. By Axiom V we may replace A by P in (12). Upon substituting from (4) we note that the resulting terms of (12) which depend on the same number ν, $\nu = j, \cdots, k$, of events have the same sign, that all $S(\nu)$, $\nu = j, \cdots, k$, occur, that no term depending on fewer than j events occurs, and that any particular $P(E_{\alpha_1}E_{\alpha_2} \cdots E_{\alpha_{j+t}})$ will occur in those of the terms of (12) the j occurring events of which are a subset of $E_{\alpha_1}, E_{\alpha_2}, \cdots, E_{\alpha_{j+t}}$ and will occur in no other term of (12). Hence the coefficient of $S(j + t)$ in (11) is $(-1)^t (j + t; t)$. This completes the proof of the theorem.

COROLLARY. If (10) is true for $\nu = j, \cdots, k$, then

$$P_{(j)} = \sum_{\nu=0}^{k-j} (-1)^{\nu} (k; j, \nu) P(E_1E_2 \cdots E_{j+\nu}). \tag{13}$$

Theorem III. The probability that at least j of the k events (1) *occur is*

$$P^{(j)} = \sum_{\nu=0}^{k-j} (-1)^{\nu} (j + \nu - 1; \nu) S(j + \nu). \tag{14}$$

Proof. If $A^{(j)}$ is the subset of Ω defined by the requirement that at least j of the events (1) occur, then $A^{(j)}$ is the sum of $k - j + 1$ disjunct sets:

$$A^{(j)} = A_{(j)} + A_{(j+1)} + \cdots + A_{(k)}. \tag{15}$$

By Axiom V we may replace A by P in (15). Substituting from (11)

$$P^{(j)} = \sum_{\nu=0}^{k-j} c_{\nu} S(j + \nu), \tag{16}$$

where

$$c_{\nu} = (j + \nu; j + \nu) - (j + \nu; 1) + \cdots + (-1)^{\nu}(j + \nu; \nu), \quad (\nu = 0, \cdots, k - j).$$

It is easy to prove that

$$(-1)^{\nu}(j + \nu - 1; \nu) = \sum_{\mu=0}^{\nu} (-1)^{\nu-\mu}(j + \nu; j + \mu). \tag{17}$$

COROLLARY. If (10) is true for $\nu = j, \cdots, k$, then

$$P^{(j)} = \sum_{\nu=0}^{k-j} (-1)^{\nu}(j + \nu - 1; \nu)(k; j + \nu)P(E_1E_2 \cdots E_{j+\nu}). \tag{18}$$

To provide examples illustrating these theorems let us consider r experiments

$$E^{(1)}, E^{(2)}, \cdots, E^{(r)} \tag{19}$$

Let $E^{(i)}$ have k mutually exclusive outcomes

$$O_1^{(i)}, O_2^{(i)}, \cdots, O_k^{(i)}. \tag{20}$$

Then, it is easy to define the spaces $\Omega^{(i)}$, $\Delta^{(i)}$ the probability function $P_i(E^{(i)})$, the combinatory product

$$\Omega = \Omega^{(1)} \times \Omega^{(2)} \times \cdots \times \Omega^{(r)},$$

the set Δ and the probability function $P(E)$ so that Axioms I, $\cdots$, V are satisfied and hence Theorems I, II, and III are valid.

We shall assume that the experiments (19) are independent.

Let

$$\bar{O}_j \qquad (j = 1, \cdots, k)$$

be the event which occurs when neither $O_j^{(1)}$ nor $O_j^{(2)}$ nor $\cdots$ nor $O_j^{(r)}$ occur. Then O_j occurs if upon performance of the experiments (19) at least one of $O_j^{(1)}, O_j^{(2)}, \cdots, O_j^{(r)}$ occur.

It is an immediate result of the definition of independence that

$$P(\bar{O}_{\alpha_1}\bar{O}_{\alpha_2}\cdots\bar{O}_{\alpha_j}) = \prod_{i=1}^{k}\{1 - P(O_{\alpha_1}^{(i)}) - \cdots - P(O_{\alpha_j}^{(i)})\}. \tag{21}$$

From Theorem I, the probability that $O_1, O_2, \cdots, O_j$ each occur while not one of $O_{j+1}, O_{j+2}, \cdots, O_k$ occurs is

$$\begin{aligned} P(O_1 \cdots O_j\bar{O}_{j+1}\cdots\bar{O}_k) &= \sum_{\nu=0}^{j}(-1)^\nu \sum_{\substack{\alpha_1,\cdots,\alpha_\nu=1\\ \alpha_1<\cdots<\alpha_\nu}}^{j} \\ &\prod_{i=1}^{r}\{1 - P(O_{j+1}^{(i)}) - \cdots - P(O_k^{(i)}) - P(O_{\alpha_1}^{(i)}) - \cdots - P(O_{\alpha_\nu}^{(i)})\}. \end{aligned} \tag{22}$$

From Theorem II, the probability that exactly j of $O_1, O_2, \cdots, O_k$ occur is

$$P_{(j)} = \sum_{\nu=0}^{j}(-1)^\nu(k-j+\nu;\nu)S(k-j+\nu), \tag{23}$$

where

$$S(k-j+\nu) = \sum_{\substack{\alpha_1,\alpha_2,\cdots,\alpha_{k-j+\nu}=1\\ \alpha_1<\alpha_2<\cdots<\alpha_{k-j+\nu}}}^{k} \prod_{i=1}^{r}\{1 - P(O_{\alpha_1}^{(i)}) - \cdots - P(O_{\alpha_{k-j+\nu}}^{(i)})\}.$$

Since the probability that at least j of $O_1, O_2, \cdots, O_k$ occur is equal to 1 minus the probability that at least $k - j + 1$ of $\bar{O}_1, \bar{O}_2, \cdots, \bar{O}_k$ occur,[12] it follows at once from Theorem III that

$$\begin{aligned} P\{\text{at least } j \text{ of } O_1, \cdots, O_k \text{ occur}\} &= \\ 1 - \sum^{j-1}(-1)^\nu(k-j+\nu;\nu)S(k-j+\nu+1). \end{aligned} \tag{24}$$

[12] There are, of course, other ways of computing these probabilities.

The case treated by Fréchet and Jordan is that which occurs when we assume $P(O_t^{(h)}) = P(O_t^{(i)})$, $(t = 1, \cdots, k)$, $(i, h = 1, \cdots, r)$ and in (24) let $j = 1$.

It is not difficult to obtain further generalizations of Jordan's distribution by defining events which occur if and only if fewer than j' of r events occur and then proceeding as above.

Certain useful generalizations of Theorems I, II, and III will now be derived.

Let the subsets of Ω

$$E_1^{(s)}, E_2^{(s)}, \cdots, E_{k^{(s)}}^{(s)} \qquad (s = 1, \cdots, p) \tag{25}$$

be elements of Δ, and let $N = k^{(1)} + k^{(2)} + \cdots + k^{(p)}$.

Let $j^{(s)} \leq k^{(s)}$, $(s = 1, \cdots, p)$; and let

$$Q^{(t)} = \prod^{t} \prod^{j^{(s)}} E_i^{(s)} \qquad (t = 1, \cdots, p), \tag{26}$$

Let

$$Q^{(t)'} = \prod_{s=1}^{t} \prod_{i=j^{(s)}+1}^{k^{(s)}} \bar{E}_i^{(s)} \qquad (t = 1, \cdots, p). \tag{27}$$

Furthermore, let for each value of s, $(s = h, \cdots, p)$, the $(k^{(s)} - j^{(s)}; \nu^{(s)})$ possible distinct selections of $\nu^{(s)}$ of the $k^{(s)} - j^{(s)}$ sets

$$E_{j^{(s)}+1}^{(s)}, E_{j^{(s)}+2}^{(s)}, \cdots, E_{k^{(s)}}^{(s)} \tag{28}$$

be arranged in some order, and, if the intersection of the $\nu^{(s)}$ sets of the $i_s{}^{\text{th}}$ selection be denoted by

$$q^{i_s}(\nu^{(s)}) \qquad (s = h, \cdots, p), \quad (i_s = 1, 2, \cdots, (k^{(s)} - j^{(s)}; \nu^{(s)})), \tag{29}$$

let

$$q^{i_h \cdots i_p}(\nu^{(h)}, \cdots, \nu^{(p)}) = \prod_{s=h}^{p} q^{i_s}(\nu^{(s)}). \tag{30}$$

There are $\prod_{s=h}^{p} (k^{(s)} - j^{(s)}; \nu^{(s)})$ sets (30), for each value of h, $(h = 1, \cdots, p)$, and any set of fixed values of $\nu^{(h)}, \cdots, \nu^{(p)}$.

Let for each value of s, $(s = h, \cdots, p)$ the $(k^{(s)}; \nu^{(s)})$ possible distinct selections of $\nu^{(s)}$ of the $k^{(s)}$ sets

$$E_i^{(s)}, \qquad (i = 1, \cdots, k^{(s)}), \tag{31}$$

be arranged in some order, and if the intersection of the sets of the $i_s{}^{\text{th}}$ selection be denoted by

$$\dot{q}^{i_s}(\nu^{(s)}) \tag{32}$$

let

$$\dot{q}^{i_h \cdots i_p}(\nu^{(h)}, \cdots, \nu^{(p)}) = \prod_{s=h}^{p} \dot{q}^{i_s}(\nu^{(s)}). \tag{33}$$

There are $\prod_{s=h}^{p} (k^{(s)}; \nu^{(s)})$ sets (33), for each value of h, $(h = 1, \cdots, p)$, and any set of fixed values of $\nu^{(h)}, \cdots, \nu^{(p)}$.

It is clear that the various sets that have been defined are elements of Δ. The fact that the sets are the events which occur if and only if certain sets of events occur is also too obvious to require further comment.

Theorem IV. The probability that of the N events (25) *the first* $j^{(s)}$ *of superscript s occur and the remaining* $k^{(s)}$ *of superscript s do not occur,* $s = 1, \cdots, p$, *is*

$$(34)\quad P(Q^{(p)} Q^{(p)\prime}) = \sum_{\nu^{(1)}=0}^{k^{(1)}-j^{(1)}} \sum_{\nu^{(2)}=0}^{k^{(2)}-j^{(2)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(1)}+\nu^{(2)}+\cdots+\nu^{(p)}} \sum_{i_1=1}^{(k^{(1)}-j^{(1)};\nu^{(1)})} \cdots \sum_{i_p=1}^{(k^{(p)}-j^{(p)};\nu^{(p)})} P[q^{i_1\cdots i_p}(\nu^{(1)} \cdots \nu^{(p)})].$$

Proof. Theorem I is a proof of Theorem IV for $p = 1$. The theorem may then be proved either by regarding it as a special case of Theorem I and collecting terms, or by induction.

COROLLARY. If, for each possible set of values of $\nu^{(1)}, \nu^{(2)}, \cdots, \nu^{(p)}$ the

$$\prod_{s=1}^{p} (k^{(s)} - j^{(s)}; \nu^{(s)})$$

terms

$$(35)\quad P[q^{i_1\cdots i_1}(\nu^{(1)}, \cdots, \nu^{(p)})]$$

are all equal, then

$$(36)\quad P(Q^{(p)} Q^{(p)\prime}) = \sum_{\nu^{(1)}=0}^{k^{(1)}-j^{(1)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(1)}+\cdots+\nu^{(p)}} \prod_{s=1}^{p} (k^{(s)} - j^{(s)}; \nu^{(s)}) P[q^{1\cdots 1}(\nu^{(1)}, \cdots, \nu^{(p)})].$$

Let, for each value of h, $(h = 1, \cdots, p)$,

$$(37)\quad S(\nu^{(h)}, \nu^{(h+1)}, \cdots, \nu^{(p)}) = \sum_{i_h=1}^{(k^{(h)};\nu^{(h)})} \cdots \sum_{i_p=1}^{(k^{(p)};\nu^{(p)})} P[Q^{(h-1)} Q^{(h-1)\prime} q^{i_h\cdots i_p}(\nu^{(h)} \cdots \nu^{(p)})].$$

It is apparent that by using (34) it is possible to obtain an expression for (37) which does not depend explicitly on $Q^{(h-1)\prime}$. In fact

$$(38)\quad S(\nu^{(h)}, \cdots, \nu^{(p)}) = \sum_{\nu^{(1)}=0}^{k^{(1)}-j^{(1)}} \cdots \sum_{\nu^{(h-1)}=0}^{k^{(h-1)}-j^{(h-1)}} (-1)^{\nu^{(1)}+\cdots+\nu^{(h-1)}} \sum_{i_1=1}^{(k^{(1)}-j^{(1)};\nu^{(1)})} \cdots \sum_{i_{h-1}=1}^{(k^{(h-1)}-j^{(h-1)};\nu^{(h-1)})} \sum_{i_h=1}^{(k^{(h)};\nu^{(h)})} \cdots \sum_{i_p=1}^{(k^{(p)};\nu^{(p)})} P[q^{i_1\cdots i_{h-1}}(\nu^{(1)}, \cdots, \nu^{(h-1)}) q^{i_h\cdots i_p}(\nu^{(h)}, \cdots, \nu^{(p)})].$$

If the different terms of (37) are all equal, then

$$(39)\quad S(\nu^{(h)}, \cdots, \nu^{(p)}) = \prod_{s=h}^{p} (k^{(s)}; \nu^{(s)}) P[Q^{(h-1)} Q^{(h-1)\prime} \dot{q}^{1,\cdots,1}(\nu^{(h)}, \cdots, \nu^{(p)})].$$

If the different terms of (38) are all equal, then

$$(40)\quad \begin{aligned} S(\nu^{(h)}, \cdots, \nu^{(p)}) = \sum_{\nu^{(1)}=0}^{k^{(1)}-j^{(1)}} \cdots \sum_{\nu^{(h-1)}=0}^{k^{(h-1)}-j^{(h-1)}} (-1)^{\nu^{(1)}+\cdots+\nu^{(h-1)}} \\ \prod_{s=1}^{h-1} (k^{(s)} - j^{(s)}; \nu^{(s)}) \prod_{s=h}^{p} (k^{(s)}; \nu^{(s)}) \\ P[q^{1\cdots 1}(\nu^{(1)}, \cdots, \nu^{(h-1)}) \dot{q}^{1\cdots 1}(\nu^{(h)}, \cdots, \nu^{(p)})]. \end{aligned}$$

Theorem V. The probability that of the N events (25) the first $j^{(s)}$ of superscript s occur and the remaining $k^{(s)}$ do not occur, $(s = 1, \cdots, h - 1)$, and exactly $j^{(s)}$ events of superscript s occur $(s = h, \cdots, p)$, is

$$(41)\quad \begin{aligned} P_{(j^{(h)}\cdots j^{(p)})}(Q^{(h-1)} Q^{(h-1)\prime}) = \sum_{\nu^{(h)}=0}^{k^{(h)}-j^{(h)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(h)}+\cdots+\nu^{(p)}} \\ \prod_{s=h}^{p} (j^{(s)} + \nu^{(s)}; \nu^{(s)}) S(j^{(h)} + \nu^{(h)}, \cdots, j^{(p)} + \nu^{(p)}). \end{aligned}$$

Proof. The theorem may be proved, either by induction using Theorem II, or by obtaining disjunct sets as in Theorem II and using Theorem IV.

Corollary I. If (39) is true for all sets of possible values of $\nu^{(h)}, \cdots, \nu^{(p)}$ then

$$(42)\quad \begin{aligned} P_{(j^{(h)}\cdots j^{(p)})}(Q^{(h-1)} Q^{(h-1)\prime}) = \sum_{\nu^{(h)}=0}^{k^{(h)}-j^{(h)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(h)}+\cdots+\nu^{(p)}} \\ \prod_{s=h}^{p} (k^{(s)}; j^{(s)}, \nu^{(s)}) P[Q^{(h-1)} Q^{(h-1)\prime} \dot{q}^{1\cdots 1}(\nu^{(h)}, \cdots, \nu^{(p)}). \end{aligned}$$

Corollary II. If (40) is true for all sets of possible values of $\nu^{(1)}, \nu^{(2)}, \cdots, \nu^{(p)}$ then

$$(43)\quad \begin{aligned} P_{(j^{(h)}\cdots j^{(p)})}(Q^{(h-1)} Q^{(h-1)\prime}) = \sum_{\nu^{(1)}=0}^{k^{(1)}-j^{(1)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(1)}+\cdots+\nu^{(p)}} \\ \prod_{s=1}^{h-1} (k^{(s)} - j^{(s)}; \nu^{(s)}) \prod_{s=h}^{p} (k^{(s)}; j^{(s)}, \nu^{(s)}) \\ P[q^{1\cdots 1}(\nu^{(1)}, \cdots, \nu^{(h-1)}) \dot{q}^{1\cdots 1}(\nu^{(h)} \cdots \nu^{(p)})]. \end{aligned}$$

Theorem VI. The probability that of the N events (25) the first $j^{(s)}$ events of superscript s occur and the remaining $k^{(s)}$ do not occur, $s = 1, \cdots, g - 1$, exactly $j^{(s)}$ events of superscript s occur $(s = g, \cdots, h - 1)$, and at least $j^{(s)}$ events of superscript s occur $(s = h, \cdots, p)$ is

$$P_{(j^{(g)}\cdots j^{(h-1)})}^{(j^{(h)}\cdots j^{(p)})}(Q^{(g-1)}Q^{(g-1)\prime}) = \sum_{\nu^{(g)}=0}^{k^{(g)}-j^{(g)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(g)}+\cdots+\nu^{(p)}}$$

$$\tag{44} \prod_{s=g}^{h-1} (j^{(s)} + \nu^{(s)}; \nu^{(s)}) \prod_{s=h}^{p} (j^{(s)} + \nu^{(s)} - 1; \nu^{(s)})$$

$$S(j^{(g)} + \nu^{(g)}, \cdots, j^{(p)} + \nu^{(p)}).$$

Proof. The theorem may be proved either by induction using Theorem III or by obtaining disjunct sets as in Theorem III and using Theorem V.

COROLLARY I. If (39) is true for all sets of possible values of

$$\nu^{(g)}, \nu^{(g+1)}, \cdots, \nu^{(p)}$$

then

$$P_{(j^{(g)}\cdots j^{(h-1)})}^{(j^{(h)}\cdots j^{(p)})}(Q^{(g-1)}Q^{(g-1)\prime}) = \sum_{\nu^{(g)}=0}^{k^{(g)}-j^{(g)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(g)}+\cdots+\nu^{(p)}}$$

$$\tag{45} \prod_{s=g}^{h-1} (k^{(s)}; j^{(s)}, \nu^{(s)}) \prod_{s=h}^{p} [(j^{(s)} + \nu^{(s)} - 1; \nu^{(s)})(k^{(s)}; j^{(s)} + \nu^{(s)})]$$

$$P[Q^{(h-1)}Q^{(h-1)\prime}\dot{q}^{1\cdots 1}(\nu^{(g)}, \cdots, \nu^{(p)})].$$

COROLLARY II. If (40) is true for all sets of possible values of $\nu^{(1)}, \nu^{(2)}, \cdots, \nu^{(p)}$ then

$$P_{(j^{(g)}\cdots j^{(h-1)})}^{(j^{(h)}\cdots j^{(p)})}(Q^{(g-1)}Q^{(g-1)\prime}) = \sum_{\nu^{(1)}=0}^{k^{(1)}-j^{(1)}} \cdots \sum_{\nu^{(p)}=0}^{k^{(p)}-j^{(p)}} (-1)^{\nu^{(1)}+\cdots+\nu^{(p)}}$$

$$\tag{46} \prod_{s=1}^{g-1} (k^{(s)} - j^{(s)}; \nu^{(s)}) \prod_{s=g}^{h-1} (k^{(s)}; j^{(s)}, \nu^{(s)}) \prod_{s=h}^{p} [(j^{(s)} + \nu^{(s)} - 1; \nu^{(s)})(k^{(s)}; j^{(s)} + \nu^{(s)})]$$

$$P[q^{1\cdots 1}(\nu^{(1)}, \cdots, \nu^{(g-1)})\,\dot{q}^{1\cdots 1}(\nu^{(g)}, \cdots, \nu^{(p)})].$$

Let us again consider the experiments (19), and let us assume that $E^{(i)}$, $(i = 1, \cdots, r)$ has as its mutually exclusive results

$$\tag{47} O_{ts}^{(i)} \qquad (t = 1, \cdots, k^{(s)});\ (s = 1, 2).$$

Let O_{ts} be the event which occurs if, upon performance of the experiments (19) at least one of the events $O_{ts}^{(1)}, O_{ts}^{(2)}, \cdots, O_{ts}^{(r)}$ occur, and let $\bar{O}_{ts}$ be the event which occurs if and only if O_{ts} does not occur.

We may state the probability that the event E_1, which occurs if and only if at least $j^{(1)}$ of the events O_{t1}, $(t = 1, \cdots, k^{(1)})$ occur, and the event E_2, which occurs if and only if at least $j^{(2)}$ of the events O_{t2}, $(t = 1, \cdots, k^{(2)})$ occur, both occur.

It is apparent that

$$\tag{48} P(E_1E_2) = 1 - P(\bar{E}_1) - P(\bar{E}_2) + P(\bar{E}_1\bar{E}_2),$$

where $\bar{E}_s$ is the event which occurs if and only if E_s does not occur, $(s = 1, 2)$.

From Theorem III

$$P(\bar{E}_s) = \sum_{\nu^{(s)}=0}^{j^{(s)}-1} (-1)^{\nu^{(s)}} (k^{(s)} - j^{(s)} + \nu^{(s)}; \nu^{(s)}) S^{(s)}(k^{(s)} - j^{(s)} + \nu^{(s)} + 1) \quad (s = 1, 2), \tag{49}$$

where

$$S^{(s)}(k^{(s)} - j^{(s)} + \nu^{(s)} + 1) = \sum_{\substack{\alpha_1, \cdots, \alpha_{k^{(s)}-j^{(s)}+\nu^{(s)}+1}=1 \\ \alpha_1 < \cdots < \alpha_{k^{(s)}-j^{(s)}+\nu^{(s)}+1}}}^{k^{(s)}} \prod_{i=1}^{r} \{1 - P(O^{(i)}_{\alpha_1 s}) - \cdots - P(O^{(i)}_{\alpha_{k^{(s)}-j^{(s)}+\nu^{(s)}+1} s})\}, \quad (s = 1, 2). \tag{50}$$

From Theorem VI

$$P(\bar{E}_1\bar{E}_2) = \sum_{\nu^{(1)}=0}^{j^{(1)}-1} \sum_{\nu^{(2)}=0}^{j^{(2)}-1} (-1)^{\nu^{(1)}+\nu^{(2)}} \prod_{s=1}^{2} (k^{(s)} - j^{(s)} + \nu^{(s)}; \nu^{(s)}) S(k^{(1)} - j^{(1)} + \nu^{(1)} + 1, k^{(2)} - j^{(2)} + \nu^{(2)} + 1), \tag{50}$$

where

$$S(k^{(1)} - j^{(1)} + \nu^{(1)} + 1, k^{(2)} - j^{(2)} + \nu^{(2)} + 1) = \sum_{i_1=1}^{(k^{(1)}; j^{(1)}-\nu^{(1)}-1)} \sum_{i_2=1}^{(k^{(2)}; j^{(2)}-\nu^{(2)}-1)} P[q^{i_1 i_2}(k^{(1)} - j^{(1)} + \nu^{(1)} + 1, k^{(2)} - j^{(2)} + \nu^{(2)} + 1)],$$

and

$$P[q^{i_1 i_2}(k^{(1)} - j^{(1)} + \nu^{(1)} + 1, k^{(2)} - j^{(2)} + \nu^{(2)} + 1)] = \prod_{i=1}^{r} \left\{1 - \sum_{\nu=1}^{k^{(1)}-j^{(1)}+\nu^{(1)}+1} P(O^{(i)}_{\alpha_\nu 1}) - \sum_{\mu=1}^{k^{(2)}-j^{(2)}+\nu^{(2)}+1} P(O^{(i)}_{\beta_\mu 2})\right\},$$

the subscripts α_ν, $(\nu = 1, \cdots, k^{(1)} - j^{(1)} + \nu^{(1)} + 1)$, being those of the i_2th selection of $k^{(1)} - j^{(1)} + \nu^{(1)} + 1$ events from $k^{(1)}$ events, and the subscripts β_μ, $(\mu = 1, \cdots, k^{(2)} - j^{(2)} + \nu^{(2)} + 1)$, being those of the i_2th selection of $k^{(2)} - j^{(2)} + \nu^{(2)} + 1$ events from $k^{(2)}$ events.

The desired probability is then obtained by substituting from (49) and (50) into (48). The procedure is perfectly general, and applies directly to situations in which $p > 2$.

We shall now investigate the results obtained by requiring that the events considered satisfy a relation of implication.

Let the subsets of Ω

$$E_{1s}, E_{2s}, \cdots, E_{ks}, \quad (s = 1, \cdots, p), \tag{51}$$

be elements of Δ, and let

$$E_{is} \subset E_{it}, \quad (i = 1, \cdots, k), \tag{52}$$

if $s < t$.

It follows that

(53) $$P(E_{is}E_{it}) = P(E_{is}), \qquad (i = 1, \cdots, k), (s < t).$$

Let $j_1 \leq j_2 \leq \cdots \leq j_t$ and let

(54) $$Q_t = \prod^{t} \prod^{j_s} E_{is}, \qquad (t = 1, 2, \cdots, p).$$

Let $j_1 \leq j_2 \leq \cdots \leq j_t$ and let

(55) $$Q'_t = \prod_{s=1} \prod_{i=j_s+1} \bar{E}_{is}, \qquad (t = 1, 2, \cdots, p).$$

From (52) and (53), it follows that

(56) $$P(Q_t Q'_t) = P\left(\left[\prod_{s=1}^{t} \prod_{i=j_{s-1}+1}^{j_s} E_{is}\right] \left[\prod_{s=1}^{t-1} \prod_{i=j_s+1}^{j_{s+1}} \bar{E}_{is}\right] \prod_{i=j_t+1}^{k} \bar{E}_{it}\right), \quad (j_0 = 0) \quad (t = 1, 2, \cdots, p).$$

Let $j_1 \leq j_2 \leq \cdots \leq j_p$ and for each value of s, $(s = 1, \cdots, p)$, consider a selection of $j_s + \nu_s$ events of second subscript s from (51). Let the p selections thus obtained be such that

$$j_s + \nu_s \leq j_{s+1}, \quad (s = 1, 2, \cdots, p), (j_{p+1} = k),$$

and if E_{is} is one of the events of the selection of events of second subscript s then the fact that $t > s$ implies that E_{it} is one of the events of the selection of events of second subscript t.

From (52) and (53), the probability of the occurrence of all the events of the p selections thus obtained is a function of $j_p + \nu_p$ events, μ_s of which are of second subscript s, $(s = 1, \cdots, p)$ where

(57) $$\mu_1 + \mu_2 + \cdots + \mu_s = j_s + \nu_s, \qquad (s = 1, \cdots, p),$$

and for a given set of values of $j_1, j_2, \cdots, j_p$ the μ_s and ν_s determine one another uniquely, $(s = 1, \cdots, p)$.

For a definite set of values of $j_1, \cdots, j_p$ and $\mu_1, \cdots, \mu_p$ or $j_1, \cdots, j_p$ and $\nu_1, \cdots, \nu_p$ there will be

$$(j_{s+1} - j_s; \nu_s) = (j_{s+1} - j_s; j_{s+1} - \mu_1 - \cdots - \mu_s), \quad (s = 1, \cdots, p), \ (j_{p+1} = k)$$

possible distinct selections of $j_s + \nu_s$, $(s = 1, \cdots, p)$ events of second subscript s, j_s of which are preassigned, from j_{s+1} events, $(s = 1, \cdots, p)$.

Let these selections be arranged in some order for each value of s, $s = 1, \cdots, p$, and let

(58) $$q_{i_1 i_2 \cdots i_p}(\mu_1, \mu_2, \cdots, \mu_p)$$

be the event which occurs when for all values of s, $(s = 1, \cdots, p)$, the events of the i_s^{th} selection of $j_s + \nu_s$ events of second subscript s all occur.[13]

[13] It is understood that the j_s preassigned events of second subscript s are among the j_t preassigned events of second subscript t, $(t > s)$ in the events (58).

A typical event (58) is

$$q_{1\cdots 1}(\mu_1, \cdots, \mu_p) = \prod_{s=1}^{p} \prod_{i=j_{s-1}+\nu_{s-1}+1}^{j_s+\nu_s} E_{is}, \qquad (j_0 + \nu_0 = 0). \tag{59}$$

There will be, for a definite j_s events of second subscript s, $(s = 1, \cdots, p)$

$$\prod_{s=1}^{p} (j_{s+1} - j_s; \nu_s), \qquad (j_{p+1} = k), \tag{60}$$

events such as (58).

For a definite set of values of $\mu_1, \cdots, \mu_p$ there will be, for each value of s, $(s = 1, \cdots, p)$

$$(k - \mu_{s-1} - \cdots - \mu_1; \mu_s), \qquad (s = 1, 2, \cdots, p)$$

possible distinct selections of $j_s + \nu_s$ events of second subscript s, $j_{s-1} + \nu_{s-1}$ of which are preassigned from k events, $(s = 1, \cdots, p)$.

Let these selections be arranged in some order for each value of s,

$$(s = 1, \cdots, p),$$

and let

$$q_{i_1 i_2 \cdots i_p}(\mu_1, \mu_2, \cdots, \mu_p) \tag{61}$$

be the event which occurs if and only if, for all values of s the events of the i_s^{th} set of $j_s + \nu_s$ events of second subscript s all occur, $(s = 1, \cdots, p)$, and the first subscripts of the events of the i_s^{th} set of events of second subscript s are among the first subscripts of the events of all the selections of events of second subscript greater than s, $(s = 1, \cdots, p)$.

There will be

$$(k; \mu_1, \mu_2, \cdots, \mu_p) \tag{62}$$

events (61) which may thus be obtained.

Theorem VII. The probability that of the pK events (51) the first j_s events of second subscript s occur and the remaining $k - j_s$ events do not occur, $s = 1, \cdots, p$, is

$$\begin{aligned} P(Q_pQ_p') = &\sum_{\nu_1=0}^{j_2-j_1} \sum_{\nu_2=0}^{j_3-j_2} \cdots \sum_{\nu_p=0}^{k-j_p} (-1)^{\nu_1+\nu_2+\cdots+\nu_p} \\ &\sum_{i_1=1}^{(j_2-j_1;\nu_1)} \sum_{i_2=1}^{(j_3-j_2;\nu_2)} \cdots \sum_{i_p=0}^{(k-j_p;\nu_p)} P[q_{i_1i_2\cdots i_p}(\mu_1, \mu_2, \cdots, \mu_p)], \end{aligned} \tag{63}$$

where the event Q_t determines the $j_s - j_{s-1} - \nu_{s-1}$ events of second subscript s, $(s = 1, \cdots, p)$, which have as first subscripts all numbers $1, 2, \cdots, j_s$ which are not among the $j_{s-1} + \nu_{s-1}$ numbers determined by the events of lower second subscript than s which are contained in $q_{i_1 \cdots i_p}(\mu_1, \cdots, \mu_p)$.

Proof. Expand (56) by means of Theorem IV.

COROLLARY. If, for each fixed set of values of $\mu_1, \mu_2, \cdots, \mu_p$ the terms (58), in number (60), are all equal, then

$$P(Q_p Q'_p) = \sum_{\nu_1=0}^{j_2-j_1} \sum_{\nu_2=0}^{j_3-j_2} \cdots \sum_{\nu_p=0}^{k-j_p} (-1)^{\nu_1+\nu_2+\cdots+\nu_p} \prod_{s=1}^{p} (j_{s+1} - j_s; \nu_s) \tag{64}$$
$$P[q_{1\cdots 1}(\mu_1, \mu_2, \cdots, \mu_p)] \qquad (j_{p+1} = k).$$

Let

$$T(\mu_1, \mu_2, \cdots, \mu_p) = \sum_{i_1=1}^{(k;\mu_1)} \sum_{i_2=1}^{(k-\mu_1;\mu_2)} \cdots \sum_{i_p=1}^{(k-\mu_1-\cdots-\mu_{p-1};\mu_p)} \tag{65}$$
$$P[\dot{q}_{i_1 i_2 \cdots i_p}(\mu_1, \mu_2, \cdots, \mu_p)].$$

If all the terms of (65) are equal, then

$$T(\mu_1, \cdots, \mu_p) = (k; \mu_1, \mu_2, \cdots, \mu_p) P[\dot{q}_{1\cdots 1}(\mu_1, \cdots, \mu_p)]. \tag{66}$$

Theorem VIII. The probability that of the pK events (51) *exactly j_s events of second subscript s, $s = 1, \cdots, p$ occur, is*

$$_{(j_1 \cdots j_p)} = \sum_{\nu_1=0}^{j_2-j_1} \sum_{\nu_2=0}^{j_3-j_2} \cdots \sum^{k-j_p} (-1)^{\nu_1+\nu_2+\cdots+\nu_p} \tag{67}$$
$$\prod^{p} (\mu_s; j_s - \mu_1 - \cdots - \mu_{s-1}) T(\mu_1, \mu_2, \cdots, \mu_p).$$

Proof. If $A_{(j_1, \ldots, j_p)}$ is the subset of Ω determined by the requirement that exactly j_s of the events (51) occur ($s = 1, \cdots, p$), then $A_{(j_1, \ldots, j_p)}$ is the sum of

$$(k; j_1, j_2 - j_1, j_3 - j_2, \cdots, j_p - j_{p-1})$$

disjunct sets which may be obtained by replacing P by A in (56) and forming (56) for all selections of $j_s - j_{s-1}$ occurring events from $k - j_{s-1}$ events, ($s = 1, \cdots, p$). By Axiom V, $P_{(j_1, \ldots, j_p)}$ is the sum of the probabilities of these disjunct sets.

Substituting from (63), it is noted that all terms (61) which depend on the same μ_s, ($s = 1, \cdots, p$), have the same sign and that all $T(\mu_1, \mu_2, \cdots, \mu_p)$ for which

$$0 \leq \nu_s \leq j_{s+1} - j_s, \qquad (s = 1, \cdots, p),$$

appear and only those appear. Furthermore any particular term (61) will occur in those of the terms (63) the $j_s - j_{s-1}$ occurring events of second subscript s, ($s = 1, \cdots, p$), of which contain a fixed ν_{s-1} events, the remaining $j_s - j_{s-1} - \nu_{s-1}$ events being a subset of the μ_s events of second subscript s, ($s = 1, \cdots, p$), that actually appear in the particular term (63). Hence the coefficient of $T(\mu_1, \cdots, \mu_p)$ is

$$(-1)^{\nu_1+\cdots+\nu_p} \prod^{p} (\mu_s; j_s - \mu_1 - \cdots - \mu_{s-1}), \qquad (\mu_0 = 0).$$

COROLLARY. If (66) is true for all sets of possible values of $\mu_1, \mu_2, \cdots, \mu_p$ then

$$P_{(j_1,\cdots,j_p)} = \sum_{\nu_1=0}^{j_2-j_1} \sum_{\nu_1=0}^{j_3-j_2} \cdots \sum_{\nu_p=0}^{k-j_p} (-1)^{\nu_1+\nu_2+\cdots+\nu_p} \tag{68}$$
$$(k; j_1, \nu_1, j_2 - j_1 - \nu_1; \nu_2, \cdots, j_p - j_{p-1} - \nu_{p-1}, \nu_p)$$
$$P[\dot{q}_{1,\cdots 1}(\mu_1, \mu_2, \cdots, \mu_p)].$$

Theorem IX. The probability that of the pk events (51) *at least* j_s, *but not more than* j_{g+1}, *events of second subscript s occur*, $(s = 1, \cdots, g)$, *and exactly* j_s *events of second subscript s occur*, $(s = g + 1, \cdots, p)$ *is*

$$P^{(j_1,\cdots,j_g)}_{(j_{g+1},\cdots,j_p)} = \sum_{\theta_2=0}^{1} \sum_{\theta_3=0}^{1} \cdots \sum_{\theta_g=0}^{1} R_{(j_{g+1}\cdots j_p)}(1, \theta_2, \cdots, \theta_g), \tag{69}$$

where, if a 1 in the i^{th} position is denoted by δ_i, $(i = 2, \cdots, g)$,

$$R_{(j_{g+1},\cdots,j_p)}(1, \delta_1, \cdots, \delta_{\gamma_1}, 0, \cdots, 0, \delta_{\gamma_2+1}, \cdots, \delta_{\gamma_3}, 0, \cdots, 0, \cdots, \delta_{\gamma_h+1}, \cdots, \delta_g)$$
$$= \sum_{\nu_p=0}^{k-j_p} \cdots \sum_{\nu_{g+1}=0}^{j_{g+2}-j_{g+1}} \sum_{\nu_g=0}^{j_{g+1}-j_g} \cdots \sum_{\nu_{\gamma_3}=j_{\gamma_4}-j_{\gamma_3}}^{j_{\gamma_4+1}-j_{\gamma_3}-1} \sum_{\nu_{\gamma_3-1}=0}^{j_{\gamma_3}-j_{\gamma_3-1}-1} \cdots \sum_{\nu_1=0}^{j_2-j_1-1} (-1)^{\nu_1+\nu_2+\cdots+\nu_p}$$
$$(j_1 + \nu_1 - 1; \nu_1) \cdots (j_{\gamma_3} + \nu_{\gamma_3} - j_{\gamma_3-1} - \nu_{\gamma_3-1} - 1; \nu_{\gamma_3}) \tag{70}$$
$$(j_{\gamma_4} + \nu_{\gamma_4} - j_{\gamma_3} - \nu_{\gamma_3} - 1; \nu_{\gamma_4}) \cdots (j_p + \nu_p - j_{p-1} - \nu_{p-1}; \nu_p)$$
$$T(j_1 + \nu_1, \cdots, j_{\gamma_3} + \nu_{\gamma_3} - j_{\gamma_3-1} - \nu_{\gamma_3-1}, 0, \cdots, 0,$$
$$j_{\gamma_4} + \nu_{\gamma_4} - j_{\gamma_3} - \nu_{\gamma_3}, \cdots, j_p + \nu_p - j_{p-1} - \nu_{p-1}).$$

Proof. We note first that there are 2^{g-1} terms in (69). Since

$$P^{(j_1,\cdots,j_g)}_{(j_{g+1},\cdots,j_p)} = \sum_{\lambda_g=j_g}^{j_{g+1}} \cdots \sum_{\lambda_2=j_2}^{\lambda_3} \sum_{\lambda_1=j_1}^{\lambda_2} P_{(\lambda_1\cdots\lambda_g\, j_{g+1}\cdots j_p)}, \tag{71}$$

the theorem may be proved by a process of repeated summation. From (67) and (71)

$$P^{(j_1)}_{(\lambda_2\cdots\lambda_g\, j_{g+1}\cdots j_p)} = \sum_{\lambda_1=j_1}^{\lambda_2} \sum_{\nu_1=0}^{\lambda_2-\lambda_1} \sum_{\nu_2=0}^{\lambda_3-\lambda_2} \cdots \sum_{\nu_p=0}^{k-j_p} (-1)^{\nu_1+\nu_2+\cdots+\nu_p} \tag{72}$$
$$(\lambda_1 + \nu_1; \nu_1)(\lambda_2 + \nu_2 - \lambda_1 - \nu_1; \nu_2) \cdots (j_p + \nu_p - j_{p-1} - \nu_{p-1}; \nu_p)$$
$$T(\lambda_1 + \nu_1, \lambda_2 + \nu_2 - \lambda_1 - \nu_1, \cdots, j_p + \nu_p - j_{p-1} - \nu_{p-1}).$$

For fixed values of $\lambda_2, \lambda_3, \cdots, \lambda_g$ there will occur in (72) all terms

$$T(j_1 + \beta_1, \lambda_2 + \nu_2 - j_1 - \beta_1, \cdots, j_p + \nu_p - j_{p-1} - \nu_{p-1}), \tag{73}$$
$$(\beta_1 = 0, \cdots, \lambda_2 - j_1), \qquad (0 \leq \nu_s \leq \lambda_{s+1} - \lambda_s), \qquad (s = 2, \cdots, p),$$
$$(\lambda_{g+s} = j_{g+s}\ s = 1, \cdots, p - g),$$

and any definite term (73) will occur in all

$$P_{(j_1+\alpha,\lambda_2,\cdots,j_p)} \tag{74}$$

for which

$$0 \leq \alpha \leq \beta_1 .$$

In (74), the definite term (73) will have coefficient

$$(75)\quad \begin{gathered}(-1)^{\beta_1-\alpha+\nu_2+\cdots+\nu_p}(j_1+\beta_1\,;j_1+\alpha)(\lambda_2+\nu_2-j_1-\beta_1\,;\nu_2)\\ \cdots(j_p+\nu_p-j_{p-1}-\nu_{p-1}\,,\nu_p),\qquad(\alpha=0,1,\cdots,\beta_1),\\ (\beta_1=0,\cdots,\lambda_2-j_1).\end{gathered}$$

Hence, in (72) the definite term (73), will have coefficient

$$(-1)^{\beta_1+\nu_2+\cdots+\nu_p}(j_1+\beta_1-1;\beta_1)(\lambda_2+\nu_2-j_1-\beta_1\,;\nu_2)\\ \cdots(j_p+\nu_p-j_{p-1}-\nu_{p-1}\,;\nu_p),$$

and

$$(76)\qquad P^{(j_1)}_{(\lambda_2,\cdots,j_p)} = R_{(\lambda_2\cdots j_p)}(1).$$

We now evaluate

$$(77)\qquad P^{(j_1 j_2)}_{(\lambda_3\cdots j_p)} = \sum_{\lambda_2=j_2}^{\lambda_3} P^{(j_1)}_{(\lambda_2\cdots j_p)}.$$

For any fixed values of $\lambda_3, \cdots, \lambda_g$, there will occur in (77) all terms

$$(78)\qquad T(j_1+\beta_1\,,j_2+\beta_2-j_1-\beta_1\,,\lambda_3+\nu_3-j_2-\beta_2\,,\\ \cdots,j_p+\nu_p-j_{p-1}-\nu_{p-1}),$$

for which either $0 \leq \beta_2 \leq \lambda_3 - j_2$; $0 \leq \beta_1 \leq j_2 - j_1 - 1$ or $\beta_1 = j_2 - j_1 + \gamma$, $0 \leq \gamma \leq \lambda_3 - j_2$; $0 \leq \beta_2 \leq \lambda_3 - j_2 - \gamma$.

Let $0 \leq \beta_1 \leq j_2 - j_1 - 1$; $0 \leq \beta_2 \leq \lambda_3 - j_2$. Then the term (78) will occur in all

$$(79)\qquad P^{(j_1)}_{(j_2+\alpha,\lambda_3,\cdots,j_p)},$$

such that

$$0 \leq \alpha \leq \beta_2 .$$

In (79), (78) will have coefficient

$$(80)\quad \begin{gathered}(-1)^{\beta_1+\beta_2-\alpha+\nu_3+\cdots+\nu_p}(j_1+\beta_1-1;\beta_1)(j_2+\beta_2-j_1-\beta_1-1;\beta_2-\alpha)\\ (\lambda_3+\nu_3-j_2-\beta_2\,;\nu_3)\cdots(j_p+\nu_p-j_{p-1}-\nu_{p-1}\,;\nu_p).\end{gathered}$$

Hence in (77), (78) will have coefficient

$$(81)\quad \begin{gathered}(-1)^{\beta_1+\beta_2+\nu_3+\cdots+\nu_p}(j_1+\beta_1-1;\beta_1)(j_2+\beta_2-j_1-\beta_1-1;\beta_2)\\ (\lambda_3+\nu_3-j_2-\beta_2\,;\nu_3)\cdots(j_p+\nu_p-j_{p-1}-\nu_{p-1}\,;\nu_p),\\ (\beta_1=0,\cdots,j_2-j_1-1),\qquad(\beta_2=0,\cdots,\lambda_3-j_2),\\ (\nu_s=0,\cdots,\lambda_{s+1}-\lambda_s),\qquad(s=3,\cdots,p);\\ (\lambda_{g+s}=j_{g+s}),\ (s=1,\cdots,p-g).\end{gathered}$$

Now let $\beta_1 = j_2 - j_1 + \gamma; 0 \leq \gamma \leq \lambda_3 - j_2 ; 0 \leq \beta_2 \leq \lambda_3 - j_2 - \gamma$. Then the term (78) will occur in all terms (79) such that

$$\gamma \leq \alpha \leq \beta_2 ,$$

and in (79), (78) will have coefficient (80). Summing for α, ($\alpha = \gamma, \cdots , \beta_2$), we obtain as the coefficient of (78) in (77)

$$0, \qquad \text{if } \beta_2 > \gamma,$$

and

$$(-1)^{\beta_1+\nu_3+\cdots+\nu_p}(j_1 + \beta_1 - 1; \beta_1)(\lambda_3 + \nu_3 - j_1 - \beta_1 ; \nu_3) \cdots (j_p + \nu_p - j_{p-1} - \nu_{p-1} ; \nu_p), \qquad \text{if } \beta_2 = \gamma.$$

Hence

$$P^{(j_1 j_2)}_{(\lambda_3 \cdots j_p)} = R_{(\lambda_3 \cdots j_p)}(1, 1) + R_{(\lambda_3 \cdots j_p)}(1, 0). \tag{82}$$

If we examine (82), we note that the result of summing with respect to λ_2 has been the replacement of (76) by two sums which are similar to (76) in that the next summation index, in this case λ_3, occurs in exactly two limits of summation. If it can be shown that the two sums which occur in (82) each result in a pair of sums after summation with respect to λ_3, or more exactly if

$$\sum_{\lambda_{s+1}=j_{s+1}}^{\lambda_{s+2}} R_{(\lambda_{s+1}, \cdots, j_p)}(1, \theta_2, \cdots , \theta_s) = R_{(\lambda_{s+2} \cdots j_p)}(1, \theta_2, \cdots , \theta_s, 1) + R_{(\lambda_{s+2}, \cdots, j_p)}(1, \theta_2, \cdots , \theta_s, 0) \tag{83}$$

then the proof will be completed.

Since the truth of (83) may be demonstrated in exactly the same way in which (82) has been shown to be true, the theorem is proved.

COROLLARY. If (66) is true for all sets of possible values of $\mu_1 , \mu_2 , \cdots , \mu_p$ then

$$\begin{aligned} &R_{(j_{g+1}, \cdots, j_p)}(1, \delta_1, \cdots \delta_{\gamma_1}, 0, \cdots , 0, \delta_{\gamma_2+1}, \cdots , \delta_{\gamma_3}, 0, \cdots 0, \cdots , \delta_{\gamma_k+1}, \cdots , \delta_g) \\ &\quad = \sum_{\nu_p=0}^{k-j_p} \cdots \sum_{\nu_{g+1}=0}^{j_{g+2}-j_{g+1}} \sum_{\nu_g=0}^{j_{g+1}-j_g} \cdots \sum_{\nu_{\gamma_3}=j_{\gamma_4}-j_{\gamma_3}}^{j_{\gamma_4+1}-j_{\gamma_3-1}} \cdots \sum_{\nu_1=0}^{j_2-j_1-1} (-1)^{\nu_1+\cdots+\nu_p} \\ &\quad (j_1 + \nu_1 - 1; \nu_1) \cdots (j_{\gamma_3} + \nu_{\gamma_3} - j_{\gamma_3-1} - \nu_{\gamma_3-1} - 1; \nu_{\gamma_3}) \\ &\quad (j_{\gamma_4} + \nu_{\gamma_4} - j_{\gamma_3} - \nu_{\gamma_3} - 1; \nu_{\gamma_4}) \cdots (j_p + \nu_p - j_{p-1} - \nu_{p-1} ; \nu_p) \\ &\quad (k; j_1 + \nu_1, \cdots , j_{\gamma_3} + \nu_{\gamma_3} - j_{\gamma_3-1} - \nu_{\gamma_3-1}, j_{\gamma_4} \\ &\qquad + \nu_{\gamma_4} - j_{\gamma_3} - \nu_{\gamma_3}, \cdots , j_p + \nu_p - j_{p-1} - \nu_{p-1}) \\ &\quad P[\dot{q}_1 \cdots _1(j_1 + \nu_1, \cdots , j_{\gamma_3} + \nu_{\gamma_3} - j_{\gamma_3-1} - \nu_{\gamma_3-1}, 0, \cdots , 0, \\ &\qquad j_{\gamma_4} + \nu_{\gamma_4} - j_{\gamma_3} - \nu_{\gamma_3}, \cdots , j_p + \nu_p - j_{p-1} - \nu_{p-1})]. \end{aligned} \tag{84}$$

Let us again consider the experiments (19) and let $E^{(i)}$ have as possible results

$$O_{j_s}^{(i)} \qquad (j = 1, \cdots, k), (s = 1, 2) \; (i = 1, 2, \cdots, r).$$

Let

$$O_{j_1}^{(i)} \supset O_{j_2}^{(i)} \qquad \begin{array}{l}(i = 1, \cdots, r),\\ (j = 1, \cdots, k),\end{array}$$

i.e. $O_{j_1}^{(i)}$ occurs whenever $O_{j_2}^{(i)}$ occurs. Furthermore let the outcomes

$$O_{11}^{(i)}, O_{21}^{(i)}, \cdots, O_{k_1}^{(i)}$$

be mutually exclusive.

Let

$$\bar{O}_{j_s}, \qquad (s = 1, 2),$$

occur if and only if none of

$$O_{j_s}^{(1)}, O_{j_s}^{(2)}, \cdots, O_{j_s}^{(k)}$$

occur.

We may wish to know the probability that at least j_1 of $\bar{O}_{11}, \cdots, \bar{O}_{k1}$ and at least j_2, $j_2 \geq j_1$, of $\bar{O}_{12}, \bar{O}_{22}, \cdots, \bar{O}_{k2}$ occur.

From Theorem IX this probability is equal to

$$P^{(j_1, j_2)} = R(1, 1) + R(1, 0), \tag{85}$$

where

$$R(1, 1) = \sum_{\nu_2=0}^{k-j_2} \sum_{\nu_1=0}^{j_2-j_1-1} (-1)^{\nu_1+\nu_2}(j_1 + \nu_1 - 1; \nu_1)$$
$$(j_2 + \nu_2 - j_1 - \nu_1 - 1; \nu_2)T(j_1 + \nu_1, j_2 + \nu_2 - j_1 - \nu_1),$$

and

$$R(1, 0) = \sum_{\nu_1=j_2-j_1}^{k-j_1} (-1)^{\nu_1}(j_1 + \nu_1 - 1; \nu_1)T(j_1 + \nu_1).$$

From (63)

$$T(j_1 + \nu_1, j_2 + \nu_2 - j_1 - \nu_1) = \sum_{i_1=1}^{(k; j_1+\nu_1)} \sum_{i_2=1}^{(k-j_1-\nu_1; j_2+\nu_2-j_1-\nu_1)} \tag{86}$$
$$P[\dot{q}_{i_1 i_2}(j_1 + \nu_1; j_2 + \nu_2 - j_1 - \nu_1)],$$

where, from (61)

$$\dot{q}_{i_1 i_2}(j_1 + \nu_1, j_2 + \nu_2 - j_1 - \nu_1) = \prod_{\nu=1}^{j_1+\nu_1} \bar{O}_{\alpha_\nu 1} \prod_{\nu=j_1+\nu_1+1}^{j_2+\nu_2} \bar{O}_{\alpha_\nu 2},$$

the subscripts

$$\alpha_1, \alpha_2, \cdots, \alpha_{j_1+\nu_1} \tag{87}$$

being the first subscripts of the i_1^{th} selection of $j_1 + \nu_1$ events of second subscript 1 from

$$\bar{O}_{11}, \bar{O}_{21}, \cdots, \bar{O}_{k1},$$

and the subscripts

$$\alpha_{j_1+\nu_1+1}, \alpha_{j_1+\nu_1+2}, \cdots, \alpha_{j_2+\nu_2},$$

being the first subscripts of the i_2^{th} selection of $j_2 + \nu_2$ events of second subscript 2, $j_1 + \nu_1$ of which are (87), from

$$\bar{O}_{12}, \bar{O}_{22}, \cdots, \bar{O}_{k2}.$$

It is easy to see that

$$P[\dot{q}_{i_1 i_2}(j_1 + \nu_1, j_2 + \nu_2 - j_1 - \nu_1)] = \prod_{i=1}^{r} \left\{1 - \sum_{\nu=1}^{j_1+\nu_1} P(O_{\alpha_\nu 1}^{(i)}) - \sum_{\nu=j_1+\nu_1+1}^{j_2+\nu_2} P(O_{\alpha_\nu 2}^{(i)})\right\}.$$

Furthermore

$$T(j_1 + \nu_1) = \sum_{i_1=1}^{(k; j_1+\nu_1)} P[\dot{q}_{i_1}(j_1 + \nu_1)], \tag{88}$$

where

$$P[\dot{q}_{i_1}(j_1 + \nu_1)] = \prod_{i=1}^{r} \left\{1 - \sum_{\mu=1}^{j_1+\nu_1} P(O_{\alpha_\mu 1}^{(i)})\right\}.$$

Substituting from (86) and (88) into (85) the desired probability is obtained.

It may be remarked that theorems which have the same relation to Theorems VII, VIII, and IX that Theorems IV, V, and VI have to Theorems I, II, and III may be obtained without much difficulty.

BIBLIOGRAPHY

(1) A. Fisher, The Mathematical Theory of Probabilities, The MacMillan Company, New York, 1930.
(2) R. A. Fisher, The Design of Experiments, Oliver and Boyd, London, 1935.
(3) M. Fréchet, Généralités sur les Probabilités, Gauthier-Villars, Paris, 1937.
(4) E. J. Gumbel, Les valeurs extrêmes des distributions statistiques, Annales de L'Institut Henri Poincaré, vol. 5, (1935), pp. 115-158.
(5) Ch. Jordan, Le théorème de probabilité de Poincaré généralisé au cas de plusieurs variables indépendantes, Acta Scientarum Mathematicarum, Szeged, vol. 7, (1934), pp. 103-111.
(6) A. Kolmogoroff, Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik, vol. 2, no. 3.
(7) H. Poincaré, Calcul des Probabilités, Gauthier-Villars, Paris, 1912.

Columbia University.

REPLY TO MR. WERTHEIMER'S PAPER

RICHMOND T. ZOCH

The attainment of rigor both in applied as well as pure mathematics is a slow process, and for this reason criticism of my paper, if constructive, is welcomed.

Properties like continuity, differentiability, and dimensionality are *local* properties, that is to say a function may be continuous or differentiable over a certain range but not outside this range, or otherwise a function may be continuous or differentiable over a given range except for singular points.

The presence of singularities in functions does not necessarily cancel their utility. Thus the function $y = \tan x$ contains points where it is discontinuous, but ordinarily it is regarded as a continuous function and the presence of these singular points seldom handicaps one when working with this function. Similarly, the function $f = \bar{x} - \frac{1}{2}\frac{\mu_3}{\mu_2}$ is a function which satisfies all four Axioms as stated in Whittaker and Robinson's book and expresses the mode of Pearson's Type III curve as a symmetric function of the measures. The fact that this function is not differentiable along the line $x_1 = x_2 = x_3 = \cdots = x_n$ will never handicap the investigator for unless the frequency distribution is clearly skew the Type III curve would not be used to represent it.

It seems that Mr. Wertheimer bases nearly all his criticisms on the tacit addition of the word "*everywhere*" to Axiom IV as stated in Whittaker and Robinson's book. The word "everywhere" is not in the statement of Axiom IV and I assumed nothing else than stated in the axiom.

If one deliberately adds the word "everywhere" to Axiom IV then nearly all my criticisms of previous writers are incorrect, unfair, and unjust. However, it does not seem that clearness and rigor in mathematics are increased by reading into an axiom a word that is not there.

Consider first the criticism in my paper which remains valid even when the word "everywhere" is added. (Schimmack uses the word "everywhere" on page 127 although Whittaker and Robinson do not.) Both Schimmack and Whittaker and Robinson proceed as at the top of page 217 of the book by the latter authors with the statement: "In this equation make $k \to 0$ then each of the quantities $\left[\frac{\partial f}{\partial x_n}\right]$ tends to a value which is independent of the x's $\cdots$." This statement rests on the tacit assumption that the quantities $\left[\frac{\partial f}{\partial x_n}\right]$ are functions of k. Even if such were true the use of tacit assumptions in a rigorous proof is objectionable, but as a matter of fact these quantities are not functions of k. Thus the particular proof given in Whittaker and Robinson's book as

well as in Schimmack's paper is altogether lacking in rigor even when the word "everywhere" is added to Axiom IV. Both Schiaparelli's and Broggi's proofs appear to be entirely rigorous if the word "everywhere" is added to Axiom IV.

In preparing my paper I assumed that no prohibition on functions which had singular points was contained in Axiom IV. In other words, I assumed since the word "everywhere" did not appear there was no valid objection to introduce and discuss functions with singularities. The functions I introduced are everywhere *continuous* but they are not *differentiable* along the line in Euclidian n-space defined by $x_1 = x_2 = x_3 = \cdots = x_n$. They are differentiable at every other point in the space.

It seems to me since Axiom IV as stated in Whittaker and Robinson's book does not exclude functions which are not everywhere differentiable that all my criticism is fair and just, and moreover nearly all my statements are correct. Mr. Wertheimer is entirely correct in pointing out that the words "everywhere" on page 181 of my paper are contradictory. As a matter of fact the whole paragraph beginning with line 7 on page 181 appears to me, on reexamining it, to be unsatisfactory. Except for this single paragraph I believe my paper to be rigorous, but I welcome further criticism.

Mr. Wertheimer's conclusions in his paragraph number 4 are clearly erroneous. To show this, consider a function of k. As $k \to 0$ any one of three situations may arise, namely: (1) The function may become infinite, (2) the function may become indeterminate, that is it may take on any value whatever, (3) the function may approach a unique finite value independent of k. Neither Schimmack nor Whittaker and Robinson nor Mr. Wertheimer has established as a definite fact that the particular type of function here in question approaches a unique finite value independent of k as $k \to 0$. The truth of the matter is that this conclusion cannot be established because the function in question does not involve k either explicitly or implicitly.

In conclusion there are two things I wish to emphasize. First, even when the word "everywhere" is added to Axiom IV, the proof given in Whittaker and Robinson's book is faulty, but if one consults the references given there in the footnotes he will find two other proofs which are rigorous with this addition to Axiom IV. Second, the mode of a skew bell shaped Pearson Frequency Curve satisfies all four axioms as stated in Whittaker and Robinson's book, and the fact that these expressions for the mode are not differentiable along a certain line is never a handicap to the statistician.

George Washington University.

CORRELATION SURFACES OF TWO OR MORE INDICES WHEN THE COMPONENTS OF THE INDICES ARE NORMALLY DISTRIBUTED

By George A. Baker

Indices are widely used in statistical analyses.[1] In many cases incorrect conclusions are drawn because indices are not uncorrelated or independent even though all of the component variables are independent. In a previous paper[2] the distribution of an index both of whose components follow the normal law was given exactly i.e. without approximation. The purpose of the present paper is to give the simultaneous distribution of two or more indices when each of the components follow the normal law. The case for two indices will be discussed in detail and the extension to more indices will be indicated.

Let x_1, x_2, and x_3, be correlated variables each being normally distributed about their respective means m_1, m_2, m_3, with standard deviations σ_1, σ_2, σ_3, and let the correlations between the variables in pairs be represented by r_{12}, r_{13}, r_{23}. Then the simultaneous distribution of these three variables will be

$$\begin{aligned}\frac{1}{(2\pi)^{\frac{3}{2}} R^{\frac{1}{2}}\sigma_1\sigma_2\sigma_3} \exp. -\frac{1}{2}\frac{1}{R}\Bigg[&\frac{R_{11}(x_1-m_1)^2}{\sigma_1^2}+\frac{R_{22}(x_2-m_2)^2}{\sigma_2^2}+\frac{R_{33}(x_3-m_3)^2}{\sigma_3^2}\\ &+2R_{12}\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2}+2R_{13}\frac{(x_1-m_1)(x_3-m_3)}{\sigma_1\sigma_3}\\ &+2R_{23}\frac{(x_2-m_2)(x_3-m_3)}{\sigma_2\sigma_3}\Bigg]dx_1\,dx_2\,dx_3\end{aligned}\tag{1}$$

where
$$R=\begin{vmatrix}1 & r_{12} & r_{13}\\ r_{12} & 1 & r_{23}\\ r_{13} & r_{23} & 1\end{vmatrix}$$

and R_{ij} are the respective second order minors of R.

[1] Rietz, H. L. "On the Frequency Distribution of Certain Ratios," Annals of Mathematical Statistics, Vol. VII, No. 3, Sept. 1936, pp. 145–153.

[2] Baker, G. A., "Distribution of the Means Divided by the Standard Deviations of Samples From Non-homogeneous Populations," Annals of Mathematical Statistics, Feb. 1932, pp. 3–5.

If we make the transformation

$$z_1 = \frac{x_1}{x_3}, \qquad x_1 = z_1 z_3$$

$$z_2 = \frac{x_2}{x_3}, \qquad x_2 = z_2 z_3$$

$$z_3 = x_3, \qquad x_3 = z_3$$

$$dx_1\, dx_2\, dx_3 = z_3^2\, dz_1\, dz_2\, dz_3$$

which is certainly valid if x_1, x_2, x_3, are all positive, then (1) becomes

$$\frac{1}{(2\pi)^{\frac{3}{2}} R^{\frac{1}{2}} \sigma_1\sigma_2\sigma_3} \exp. -\frac{1}{2}\frac{1}{R}\left[\frac{R_{11}(z_1z_3 - m_1)^2}{\sigma_1^2} + \frac{R_{22}(z_2z_3 - m_2)^2}{\sigma_2^2}\right.$$

$$(2) \qquad + \frac{R_{33}(z_3 - m_3)^2}{\sigma_3^2} + 2R_{12}\frac{(z_1z_3 - m_1)(z_2z_3 - m_2)}{\sigma_1\sigma_2} + 2R_{13}\frac{(z_1z_3 - m_1)(z_3 - m_3)}{\sigma_1\sigma_3}$$

$$\left. + 2R_{23}\frac{(z_2z_3 - m_2)(z_3 - m_3)}{\sigma_2\sigma_3}\right] z_3^2\, dz_1\, dz_2\, dz_3 .$$

If x_1, x_2, x_3 are all positive the corresponding distribution of z_1 and z_2 can be obtained by integrating (2) between the limits 0 and ∞ with respect to z_3. If x_1, x_2, x_3 are all negative z_1 and z_2 are again both positive so that in order to get the total distribution for z_1 and z_2 it is necessary to add to the integral of (2) between the limits 0 and ∞ with respect to z_3 the similar integral of (2) with z_3 replaced by $-z_3$. The result is

$$(3) \qquad \frac{2e^{-\frac{1}{2}\frac{c}{R}} e^{\frac{1}{2}\frac{b^2}{Ra}}}{(2\pi)^{\frac{3}{2}} R^{\frac{1}{2}} \sigma_1\sigma_2\sigma_3}\left[\frac{\sqrt{\pi}}{\sqrt{2}}\frac{R^{\frac{3}{2}}}{a^{\frac{3}{2}}} - \frac{b^2}{a^2}\int_0^{\frac{b}{\sqrt{R}\sqrt{a}}} e^{-\frac{1}{2}z^2}\, dz + \frac{R^{\frac{1}{2}} b^2}{a^{\frac{5}{2}}}\frac{\sqrt{\pi}}{\sqrt{2}}\right]$$

where

$$a = \frac{R_{11}}{\sigma_1^2} z_1^2 + \frac{R_{22}}{\sigma_2^2} z_2^2 + \frac{R_{33}}{\sigma_3^2} + \frac{2R_{12}}{\sigma_1\sigma_2} z_1 z_2 + \frac{2R_{13}}{\sigma_1\sigma_3} z_1 + \frac{2R_{23}}{\sigma_2\sigma_3} z_2$$

$$b = \frac{R_{11}}{\sigma_1^2} m_1 z_1 + \frac{R_{22}}{\sigma_2^2} m_2 z_2 + \frac{R_{33}}{\sigma_3^2} m_3 + \frac{R_{12}}{\sigma_1\sigma_2} z_1 m_2 + \frac{R_{12}}{\sigma_1\sigma_2} m_1 z_2 + \frac{R_{13}}{\sigma_1\sigma_3} m_3 z_1$$

$$+ \frac{R_{13}}{\sigma_1\sigma_3} m_1 + \frac{R_{23}}{\sigma_2\sigma_3} m_3 z_2 + \frac{R_{23}}{\sigma_2\sigma_3} m_2$$

$$c = \frac{R_{11}}{\sigma_1^2} m_1^2 + \frac{R_{22}}{\sigma_2^2} m_2^2 + \frac{R_{33}}{\sigma_3^2} m_3^2 + \frac{2R_{12}}{\sigma_1\sigma_2} m_1 m_2 + \frac{2R_{13}}{\sigma_1\sigma_3} m_1 m_3 + \frac{2R_{23}}{\sigma_2\sigma_3} m_2 m_3 .$$

The same result (3) is obtained for z_1, and z_2 negative, z_1 positive and z_2 negative, z_1 negative and z_2 positive. That is (3) is the simultaneous distribution of z_1 and z_2. The extension to more than 2 indices is immediate. The form of the distribution of the indices and the denominator variable is the same as (2)

except that a, b, and c, the coefficients of z_3^2, z_3 and the constant term respectively in the exponent of e, will be different in that they will include the new indices and the exponent on the denominator variable will be the same as the number of indices involved. The distribution of the indices will again be obtained by integrating from 0 to ∞ with respect to the denominator variable.

The case when all of the variables x_1, x_2, x_3 are independent is especially interesting. If r_{12}, r_{13}, r_{23} are all zero then $R = R_{11} = R_{22} = R_{33} = 1$, $R_{12} = R_{13} = R_{23} = 0$ and a, b, c, become a', b', c', respectively.

$$a' = \frac{z_1^2}{\sigma_1^2} + \frac{z_2^2}{\sigma_2^2} + \frac{1}{\sigma_3^2}$$

$$b' = \frac{m_1 z_1}{\sigma_1^2} + \frac{m_2 z_2}{\sigma_2^2} + \frac{m_3}{\sigma_3^2}$$

$$c' = \frac{m_1^2}{\sigma_1^2} + \frac{m_2^2}{\sigma_2^2}\frac{m_3^2}{\sigma_3^2}$$

Under these conditions and the further condition that m_1, m_2, m_3 are large with respect to σ_1, σ_2, σ_3 respectively so that the integral term of (3) maybe neglected (3) becomes

$$(4)\qquad \frac{e^{-\frac{1}{2}\left(\frac{m_1^2}{\sigma_1^2}+\frac{m_2^2}{\sigma_2^2}+\frac{m_3^2}{\sigma_3^2}\right)}}{2\pi\sigma_1\sigma_2\sigma_3} \frac{e^{\frac{1}{2}\frac{\left(\frac{m_1 z_1}{\sigma_1^2}+\frac{m_2 z_2}{\sigma_2^2}+\frac{m_3}{\sigma_3^2}\right)^2}{\left(\frac{z_1^2}{\sigma_1^2}+\frac{z_2^2}{\sigma_2^2}+\frac{1}{\sigma_3^2}\right)}}}{\left(\frac{z_1^2}{\sigma_1^2}+\frac{z_2^2}{\sigma_2^2}+\frac{1}{\sigma_3^2}\right)^{\frac{3}{2}}} \Bigg| \, 1 + \frac{\left(\frac{m_1 z_1}{\sigma_1^2}+\frac{m_2 z_2}{\sigma_2^2}+\frac{m_3}{\sigma_3^2}\right)^2}{\left(\frac{z_1^2}{\sigma_1^2}+\frac{z_2^2}{\sigma_2^2}+\frac{1}{\sigma_3^2}\right)}$$

It is clear that z_1 and z_2 are not independent in the probability sense for distribution (4).

The question as to the possibility of having the variables independent and the indices independent at the same time arises. Denote the distribution functions of x_1, x_2, x_3, by $X_1(x_1)$, $X_2(x_2)$, $X_3(x_3)$ and of z_1, z_2 by $Z_1(z_1)$, $Z_2(z_2)$. Then, if $x_i \geq 0$, $i = 1, 2, 3$ it is necessary that

$$(5)\qquad \int_a^b X_1(z_3 z_1) X_2(z_3 z_2) X_3(z_3) z_3^2 \, dz_3 = Z_1(z_1) Z_2(z_2)$$

a and b being suitable limits.

For instance, let

$$X_1(x_1) = \frac{1}{x_1^2}, \qquad 1 \leq x_1 \leq 3$$

$$X_2(x_2) = \frac{1}{x_2^2}, \qquad 1 \leq x_2 \leq 3$$

$$X_3(x_3^2) = x_3^2, \qquad 1 \leq x_3 \leq 2$$

then

$$Z_1(z_1) = \frac{c_1}{z_1^2}$$

$$Z_2(z_2) = \frac{c_2}{z_2^2}$$

for value of z_1 and z_2 within a straight line sided area the corners of which are $(\frac{1}{2}, \frac{1}{2})$, $(\frac{1}{2}, 1)$, $(1, 1)$ and $(1, 2)$. z_1, and z_2 are not uncorrelated throughout their entire set of values but are for this particular set of values. Thus is appears that it is possible that the indices may be independent when the variables are, but not necessarily so.

Indices should be used with care since it is very easy to draw invalid conclusions from the consideration of them. Usually it is better to use partial correlation analysis to remove the influence of a third factor than to calculate indices.

THE TYPE B GRAM-CHARLIER SERIES

By Leo A. Aroian

While much attention has been devoted to the Type A Gram-Charlier series for the graduation of frequency curves, the Type B series has been somewhat neglected. However the numerical examples to be presented later will show that the Type B series is very useful for the graduation of skew frequency curves. Wicksell[1] has demonstrated that the Gram-Charlier series may be developed from the same law of probability which forms the basis of the Pearson system of frequency curves. Rietz[2] following Wicksell gives a derivation of the Gram-Charlier series based on the binomial $(q + p)^n$. Jordan[3] gives a method for fitting Type B based on certain orthogonal polynomials which he calls G. He uses factorial moments because of the resulting ease in finding the values of the constants.

We shall consider the Type B series for a distribution of equally distanced ordinates at non-negative values of x. We shall find the values of the first few terms of the series and shall also show how the values of later coefficients may easily be found. We write the Type B series in the form

$$(1)\quad F(x) = c_0 + c_1\Delta\psi(x) + c_2\Delta^2\psi(x) + c_3\Delta^3\psi(x) + c_4\Delta^4\psi(x) + c_5\Delta^5\psi(x) + c_6\Delta^6\psi(x)$$

where

$$(2)\quad \begin{aligned} \psi(x) &= \frac{e^{-m}m^x}{x!}, \qquad m = \mu_1', \text{ the mean,} \\ \Delta\psi(x) &= \psi(x) - \psi(x-1) \quad \text{for } x = 0, 1, 2, \cdots s. \end{aligned}$$

Let $f(x)$ give the ordinates of the observed distribution of relative frequencies, so that $\Sigma f(x) = 1$. To determine the coefficients $c_0, c_1, c_2, \cdots, c_6$, we have, using the method of moments,

$$(3)\quad \begin{aligned} &\Sigma[c_0\psi(x) + c_1\Delta\psi(x) + c_2\Delta^2\psi(x) + c_3\Delta^3\psi(x) + \ldots + c_6\Delta^6\psi(x)] = \Sigma f(x) = 1. \\ &\Sigma x[c_0\psi(x) + c_1\Delta\psi(x) + \ldots\ldots\ldots + c_6\Delta^6\psi(x)] = \Sigma x f(x) = m. \\ &\Sigma x^2[c_0\psi(x) + c_1\Delta\psi(x) + \ldots\ldots\ldots + c_6\Delta^6\psi(x)] = \Sigma x^2 f(x) = \mu_2'. \\ &\Sigma x^3[c_0\psi(x) + \ldots\ldots\ldots + c_6\Delta^6\psi(x)] = \Sigma x^3 f(x) = \mu_3'. \\ &\Sigma x^4[c_0\psi(x) + \ldots\ldots\ldots + c_6\Delta^6\psi(x)] = \Sigma x^4 f(x) = \mu_4'. \\ &\Sigma x^5[c_0\psi(x) + \ldots\ldots\ldots + c_6\Delta^6\psi(x)] = \Sigma x^5 f(x) = \mu_5'. \\ &\Sigma x^6[c_0\psi(x) + \ldots\ldots\ldots + c_6\Delta^6\psi(x)] = \Sigma x^6 f(x) = \mu_6'. \end{aligned}$$

Hence we must find the values of

(4) $$\sum_{x=0}^{x=s} x^n \Delta^p \psi(x), \qquad \begin{array}{l} n = 0, 1, 2, 3 \cdots \\ p = 0, 1, 2, 3 \cdots \end{array}$$

defining $\Delta^0\psi(x) = \psi(x)$. We assume that we are dealing with distributions in which s is large, and that the error involved in substituting $\sum_{x=0}^{\infty} x^n \Delta^p \psi(x)$ for $\sum_{x=0}^{s} x^n \Delta^p \psi(x)$ is negligible. To find these summations in a straightforward manner would involve too much labor, so we shall briefly discuss some properties of the generating function, $\psi(x) = \dfrac{e^{-m} m^x}{x!}$, the Poisson exponential, very useful in the graduation of frequency distributions of rare events. The first eight moments about the origin are:

(5)
$$\begin{aligned}
\mu_0' &= 1 = \Sigma\psi(x), \qquad \mu_1' = m = \Sigma x\psi(x), \qquad \mu_2' = m + m^2 = \Sigma x^2\psi(x) \\
\mu_3' &= m + 3m^2 + m^3 = \Sigma x^3\psi(x) \\
\mu_4' &= m + 7m^2 + 6m^3 + m^4 = \Sigma x^4\psi(x) \\
\mu_5' &= m + 15\, m^2 + 25m^3 + 10m^3 + m^5 = \Sigma x^5\psi(x) \\
\mu_6' &= m + 31m^2 + 90m^3 + 65m^4 + 15m^5 + m^6 = \Sigma x^6\psi(x) \\
\mu_7' &= m + 63m^2 + 301m^3 + 350\, m^4 + 140m^5 + 21m^6 + m^7 = \Sigma x^7\psi(x) \\
\mu_8' &= m + 127m^2 + 966m^3 + 1701m^4 + 1050m^5 + 256m^6 + 28m^7 + m^8 \\
&= \Sigma x^8\psi(x)
\end{aligned}$$

These may be found by the formula given by Jordan,[3]

(6) $$\mu_{s+1}' = m\left(\mu_s' + \frac{d\mu_s'}{dm}\right).$$

Proof: $$\frac{d\psi(x)}{dm} = \frac{x\psi(x)}{m} - \psi(x).$$

We multiply by x^n and sum, giving (6). This result may readily be proved also by means of recursion formulas without differentiation. Now we must find the values of

$$\sum_0^{\infty} x^n \Delta^p \psi(x) \qquad \begin{array}{l} n = 0, 1, 2, \\ p = 1, 2, 3, \end{array}$$

We do this by proving

(7) $$\sum_{x=0} x^n \Delta^{s+1} \psi(x) = -\frac{d}{dm} \sum_{x=0} x^n \Delta^s \psi(x).$$

Now

$$\frac{d\psi(x)}{dm} = \psi(x-1) - \psi(x) = -\Delta\psi(x). \tag{8}$$

Hence

$$\frac{d}{dm}\Delta^s\psi(x) = \frac{d}{dm}\Big[\psi(x) - \binom{s}{1}\psi(x-1) + \binom{s}{2}\psi(x-2) + \cdots + (-1)^s\psi(x-s)\Big],$$

since $\Delta^s\psi(x) = \psi(x) - \binom{s}{1}\psi(x-1) + \binom{s}{2}\psi(x-2) + \cdots + (-1)^s\psi(x-s).$

Then by (8)

$$\frac{d}{dm}\Delta^s\psi(x) = \Big[\psi(x-1) - \psi(x) - \binom{s}{1}\psi(x-2) + \binom{s}{1}\psi(x-1) + \binom{s}{2}\psi(x-3) - \binom{s}{2}\psi(x-2) + \cdots + (-1)^s\psi(x-s-1) - (-1)^s\psi(x-s)\Big].$$

$$\begin{aligned} \frac{d}{dm}\Delta^s\psi(x) &= -\psi(x) + \binom{s+1}{1}\psi(x-1) - \binom{s+1}{2}\psi(x-2) + \cdots - (-1)^s\psi(x-s-1). \\ &= -\Big[\psi(x) - \binom{s+1}{1}\psi(x-1) + \binom{s+1}{2}\psi(x-2) + \cdots + (-1)^s\psi(x-s-1)\Big]. \\ &= -\Delta^{s+1}\psi(x). \end{aligned} \tag{9}$$

We multiply (9) by x^n, sum with respect to x, giving (7).

Thus by use of (7) and (5) we get:

$$\begin{aligned} \Sigma\Delta^p\psi(x) &= 0, \quad p = 1, 2, 3, \cdots \\ \Sigma x\Delta\psi(x) &= -\frac{dm}{dm} = -1. \\ \Sigma x^2\Delta\psi(x) &= -\frac{d}{dm}\Sigma x^2\psi(x) = -\frac{d}{dm}(m + m^2) = -2m - 1. \\ \Sigma x^3\Delta\psi(x) &= -3m^2 - 6m - 1. \\ \Sigma x^4\Delta\psi(x) &= -4m^3 - 18m^2 - 14m - 1. \\ \Sigma x^5\Delta\psi(x) &= -5m^4 - 40m^3 - 75m^2 - 30m - 1. \end{aligned} \tag{10}$$

$$\Sigma x^6\Delta\psi(x) = -6m^5 - 75m^4 - 260m^3 - 270m^2 - 62m - 1.$$
$$\Sigma x\Delta^2\psi(x) = 0, \quad \Sigma x^2\Delta^2\psi(x) = 2, \quad \Sigma x^3\Delta^2\psi(x) = 6m + 6.$$
$$\Sigma x^4\Delta^2\psi(x) = 12m^2 + 36m + 14.$$
$$\Sigma x^5\Delta^2\psi(x) = 20m^3 + 120m^2 + 150m + 30.$$
$$\Sigma x^6\Delta^2\psi(x) = 30m^4 + 300m^3 + 780\, m^2 + 540m + 62.$$
$$\Sigma x\Delta^3\psi(x) = 0, \quad \Sigma x^2\Delta^3\psi(x) = 0, \quad \Sigma x^3\Delta^3\psi(x) = -6.$$
$$\Sigma x^4\Delta^3\psi(x) = -24m - 36, \quad \Sigma x^5\Delta^3\psi(x) = -60m^2 - 240m - 150.$$
$$\Sigma x^6\Delta^3\psi(x) = -120m^3 - 900m^2 - 1560m - 540.$$
$$(10) \quad \Sigma x\Delta^4\psi(x) = 0, \quad \Sigma x^2\Delta^4\psi(x) = 0, \quad \Sigma x^4\Delta^4\psi(x) = 24.$$
$$\Sigma x^5\Delta^4\psi(x) = 120m + 240, \quad \Sigma x^3\Delta^4\psi(x) = 0.$$
$$\Sigma x^6\Delta^4\psi(x) = 360m^2 + 1800m + 1560.$$
$$\Sigma x\Delta^5\psi(x) = 0, \quad \Sigma x\Delta^6\psi(x) = 0.$$
$$\Sigma x^2\Delta^5\psi(x) = 0, \quad \Sigma x^2\Delta^6\psi(x) = 0.$$
$$\Sigma x^3\Delta^5\psi(x) = 0, \quad \Sigma x^3\Delta^6\psi(x) = 0.$$
$$\Sigma x^4\Delta^5\psi(x) = 0, \quad \Sigma x^4\Delta^6\psi(x) = 0.$$
$$\Sigma x^5\Delta^5\psi(x) = -120, \quad \Sigma x^5\Delta^6\psi(x) = 0.$$
$$\Sigma x^6\Delta^5\psi(x) = -720m - 1800, \quad \Sigma x^6\Delta^6\psi(x) = 720.$$

Finally we substitute from (5) and (10) into (3), and for μ_n' we substitute $\mu_n' = \sum_{r=0}^{n} \binom{n}{r} \mu_{n-r} m^r$. Hence

$$c_0 = 1$$
$$c_1 = 0$$
$$c_2 = \tfrac{1}{2}(\mu_2 - m).$$
$$(11) \quad c_3 = -\frac{1}{3!}(\mu_3 - 3\mu_2 + 2m).$$
$$c_4 = \frac{1}{4!}[\mu_4 - 6\mu_3 + \mu_2(11 - 6m) + 3m(m - 2)].$$
$$c_5 = -\frac{1}{5!}[\mu_5 - 10\mu_4 - \mu_3(10m - 25) + 50\mu_2(m - 1) - 4m(5m - 6)].$$
$$c_6 = \frac{1}{6!}[\mu_6 - 15\mu_5 + \mu_4(85 - 15m) + \mu_3(130m - 225) + \mu_2(45m^2 - 375m + 274) - 15m^3 + 130m^2 - 120m].$$

It may be asked whether criteria may be given as guides for the use of Type B. In general Type B may be tried if either the skewness of the distribution to be

fitted is considerable, $\alpha_3 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} > .6$, or if $m = \mu_2 = \mu_3$ approximately. The latter condition strictly would mean that $\psi(x)$ alone is sufficient for a good graduation, if the fourth moment, μ_4, is not used. The examples which follow are arranged to facilitate comparison with the Pearson system of frequency curves. We have an example each of Type I, III, IV, V, VI, and an example of the normal curve.

Type I. Table 1. Here $\alpha_3 > .6$ although $m \neq \mu_2 \neq \mu_3$. The first four moments, unadjusted, give an excellent fit by Type B, which is not quite as good as Type I. The degrees of freedom, according to Fisher,[4] have been taken into consideration here in applying the x^2 test. The two classes 13, 14, were grouped together for the x^2 test. The actual numerical work is easily done on a calculating machine, although logarithms are necessary to find the value of e^{-m}. This example and the remaining are all taken from Elderton[5] with the exception of Type IV which is from A. Fisher.[6]

Type III. Table 2. The unadjusted moments are used. Here $\alpha_3 = 2.0833 > .6$, and $m = \mu_2$ approximately. The fit by Type B is slightly better than that by Type III. We have for Type III $P(x^2 \geq 12.8) = .007$, $n = 3$, while for Type B, $P(x^2) \geq 9.4 = .025$ $n = 3$. Moreover the standard error of prediction for Type III is 11.2 and for Type B is 7.7.

Type IV. Table 3. The rough moments were used. Although $\alpha_3 = .48 < .6$, Type B gives a fine fit since $m = \mu_2 = \mu_3$ approximately. Here the results are given for Type B using 2, 3, and 4 terms of the series. This was done to show how the distribution changes with the addition of more terms. The superiority of Type B over Type IV is evident. The results for Type IV are taken from the class notes of Professor C. C. Craig.

Type V. Table 4. Using the adjusted moments we have a comparison among Types V, A, and B. While the graduations may seem satisfactory, the x^2 test shows that the fit is poor in each case. The order of merit is Type V, Type B, and then Type A. The negative frequencies which appear in Type B may be due to the use of the adjusted moments. If we use the rough moments, the negative frequencies disappear. On the whole the fit by means of the adjusted moments is superior.

Type VI. Table 5. Type VI using the adjusted moments gives an excellent fit. Even though α_3 is considerable, and $\mu_2 = \mu_3$ approximately, four moments with Type B give a poor fit, and five moments, adjusted, achieve a very small gain. Five moments using the unadjusted moments give some improvement, but the -2 frequency in the first class is objectionable.

Normal Curve. Table 6. The normal curve provides a fine fit. $P(x^2 \geq .9) = .96$, $n = 6$. The first two and the last two classes were grouped together for the test. The fit by Type B is less probable, $P(x^2 \geq 8) = .15$, $n = 5$. Type B has two discrepancies, the negative frequencies, and the fact that the total frequencies (neglecting the -1) is 352. That Type B does so well is in itself quite amazing!

TABLE 1

x	Actual frequency	Frequency computed by Pearson Type I	Frequency given by Type B
0	34	44	42.4
1	145	137	121.3
2	156	149	168.7
3	145	142	156.8
4	123	127	120.5
5	103	108	94.9
6	86	88	82.9
7	71	69	72.2
8	55	51	56.7
9	37	36	38.0
10	21	24	23.1
11	13	14	12.0
12	7	7	5.7
13	3	3	2.4
14	1	1	.9

$m = 4.175$ $\quad\alpha_3 = .712247\quad$ Type I $P(\chi^2 \geq 4.36) = .88$

$\mu_2 = 7.66237$ $\quad\alpha_4 = 2.95214\quad$ n (number of degrees of freedom) $= 9$

$\mu_3 = 15.1069$ $\quad c_2 = 1.74368$

$\mu_4 = 173.326$ $\quad c_3 = -.078298\quad$ Type B $P(\chi^2 \geq 9.67) = .37$

$c_4 = +.094592\quad n = 9$

$$F(x) = \psi(x) + 1.74368\ \Delta^2\psi(x) - .078298\ \Delta^3\psi(x) + .094592\ \Delta^4\psi(x).$$

TABLE 2

x	Actual frequency	Frequency computed by Type III	Fequency by Type B
0	44	59	48.1
1	135	111	121.6
2	45	45	58.5
3	12	20	10.4
4	8	9	3.5
5	3	4	4.3
6	1	2	2.9
7	3	1	1.2

$m = 1.33466$ $\quad \alpha_3 = \dfrac{\mu_3}{\mu_2^{3/2}} = 2.0833 \quad c_2 = .05356$

$\mu_2 = 1.44179$ $\quad c_3 = -.32510$

$\mu_3 = 3.60662$

$$F(x) = \psi(x) + .05356\Delta^2\psi(x) - .32510\Delta^3\psi(x)$$

TABLE 3

Number of alpha particles from a bar of polonium in intervals of $\frac{1}{8}$ of one minute

x	Frequency	Type IV	Type B 2 terms	Type B 3 terms	Type B 4 terms
0	57	50	49.5	49.0	58.2
1	203	183	201.3	201.0	199.8
2	383	392	403.4	404.3	386.1
3	525	544	532.3	533.8	523.9
4	532	539	520.6	521.5	532.1
5	408	417	402.6	402.5	418.2
6	273	250	254.8	254.4	260.2
7	139	131	137.1	136.7	134.0
8	45	61	64.0	63.9	56.7
9	27	26	26.1	26.2	22.9
10	10	12	9.4	9.6	8.6
11	4	4	3.0	3.1	3.6
12	0	1	.9	.9	1.6
13	1	0	.2	.2	.8
14	1	0	.0	.0	.3

$m = 3.87155 \qquad \alpha_3 = .47844$

$\mu_2 = 3.69477 \qquad \alpha_4 = 3.506536$

$\mu_3 = 3.39791$

$\mu_4 = 47.86888$

$$F(x) = \psi(x) - .08839\Delta^2\psi(x) - .00930\Delta^3\psi(x) + .16810\Delta^4\psi(x).$$

Type B, 4 terms $P(\chi^2 \geq 4.50) = .72, n = 7$

Type IV $P(\chi^2 \geq 10.8) = .15, n = 7$

TABLE 4

Mortality Among Female Nominees

x	Deaths	Elderton Type V	Type A	Type B 2 terms	Type B 3 terms	Type B 5 terms	Type B 5 terms
0	4	4	2	1.4	−6.9	−.4	4.1
1	18	10	15	26.3	7.1	9.4	13.1
2	53	80	78	109.7	100.1	84.6	77.4
3	265	261	235	248.3	268.4	252.3	242.5
4	438	441	426	379.5	418.8	425.9	427.4
5	525	480	521	432.7	461.0	484.0	494.1
6	342	381	411	388.8	388.4	402.6	408.1
7	253	247	225	285.4	263.5	259.0	253.9
8	128	137	107	170.8	145.5	132.2	124.9
9	82	68	66	84.3	68.3	58.6	54.1
10	28	32	44	32.9	28.2	26.2	26.4
11	12	14	22	8.6	11.0	13.9	16.4
12	8	6	8	−.01	4.7	8.2	10.7
13	5	3	2	−2.1	2.1	4.3	5.9
14	1	1	0	−1.5	1.3	2.0	2.5

Adjusted moments:

$m = 5.30435$ $\alpha_3 = .703564$

$\mu_2 = 3.573345$ $\alpha_4 = 3.996196$

$\mu_3 = +4.752437$

$\mu_4 = 51.02659$

$\mu_5 = 193.439125$

Rough moments:

$m = 5.30435$

$v_2 = 3.65668$

$v_3 = 4.752437$

$v_4 = 52.85276$

$v_5 = 197.39949$

Type A: $f(t) = \varphi(t) + .117261\,\varphi^3(t) + .041508\varphi^4(t)$

Type B: $F(x) = \psi(x) - .86550\Delta^2\psi(x) - .77352\Delta^3\psi(x)$
$+ .02814\Delta^4\psi(x) + .57459\Delta^5\psi(x)$

Using uncorrected moments

Type B: $F(x) = \psi(x) - .82384\Delta^2\psi(x) - .73185\Delta^3\psi(x)$
$+ .03192\Delta^4\psi(x) + .94033\Delta^5\psi(x)$
(last column above)

TABLE 5

x	Frequency	Type VI	Type B 4 terms	Type B 5 terms
0	1	1	−9.5	−2.0
1	56	50	83.2	69.9
2	167	168	141.6	143.1
3	98	100	102.3	110.7
4	34	36	41.5	40.2
5	9	10	8.7	4.6
6	2	2	.05	2.0
7	1	.5	−.4	1.0

Corrected moments:	Rough moments:
$m = 2.402174$	$m = 2.402174$
$\mu_2 = .928835$	$\mu_2 = 1.012169$
$\mu_3 = .893096$	$\mu_3 = .893096$
$\mu_4 = 4.088800$	$\mu_4 = 4.313176$
	$\mu_5 = 11.28304$
	$\alpha_3 = .87704$
	$\alpha_4 = 4.2101$

Type B, adjusted moments:

$$F(x) = \psi(x) - .73667\Delta^2\psi(x) - .48516\Delta^3\psi(x) - .06424\Delta^4\psi(x) + .10365\Delta^5\psi(x)$$

*Type B, rough moments:

$$F(x) = \psi(x) - .69805\Delta^2\psi(x) - .44654\Delta^3\psi(x) - .06587\Delta^4\psi(x) + .15165\Delta^5\psi(x)$$

* This is used in last column of above. There is a slight error here, which however will not affect the results materially. The third decimal place may be slightly wrong.

TABLE 6

Normal curve

x	Frequency	Normal curve	Type B
0	.6	.6	2.3
1	2.8	2.7	4.7
2	11.5	10.9	8.7
3	27.7	30.1	25.2
4	59.1	58.4	55.2
5	84.7	80.1	79.5
6	74.1	76.9	80.1
7	50.5	52.2	58.1
8	23.2	25.0	29.7
9	12.2	8.4	8.6
10	1.3	2.4	−.9

Moments corrected:

$m = 5.393443$

$\mu_2 = 2.769635$

$\mu_3 = .029805,\ \mu_4 = 22.40663$

$\alpha_3 = .0064$

$\alpha_4 = 2.920997$

Type B: $F(x) = \psi(x) - 1.3119\Delta^2\psi(x) - .4179\Delta^3\psi(x) + 2.1625\Delta^4\psi(x)$

Colorado State College

REFERENCES

(1) Wicksell, S. D.: Contributions to the analytical theory of sampling, Archiv för Matematik, Astronomi och Fysik, Vol. 17 no. 19 (1923) pp. (1–46)
(2) Rietz, H. L.: Mathematical Statistics. pp. (156–172).
(3) Jordan, C.: Statistique Mathématique pp. 36–40, 90–92, 98–102, 240, 282–283.
(4) Fisher, R. A.: Statistical Methods for Research Workers. (5th edition.)
(5) Elderton, W. P.: Frequency Curves and Correlation.
(6) Fisher, A.: The Mathematical Theory of Probabilities.

A TEST OF A SAMPLE VARIANCE BASED ON BOTH TAIL ENDS OF THE DISTRIBUTION

By John W. Fertig

With the assistance of Elizabeth A. Proehl[1]

(1) Introduction

In testing the hypothesis, say H_0, that an observed sample E of size N has been drawn from a normal population for which the standard deviation, σ, has a particular value, σ_0, one may form the ratio

$$v = \underset{i=1}{\overset{N}{S}} (x_i - m)^2/\sigma_0^2 = \frac{Nd^2}{\sigma_0^2} \qquad \text{(I)}$$

if the population mean m be known, or

$$v' = \underset{i=1}{\overset{N}{S}} (x_i - \bar{x})^2/\sigma_0^2 = \frac{Ns^2}{\sigma_0^2} \qquad \text{(II)}$$

where $\bar{x}$ is the sample mean, if the population mean be unknown. The probability of obtaining a larger (or smaller) value of v or v' than that observed may readily be obtained from the appropriate tail area of the χ^2 distribution with $n = N$ or $n = (N - 1)$ degrees of freedom respectively. The alternative hypotheses to H_0 concerning the normal populations from which the sample may have been drawn assign different values to σ and form a set of hypotheses, Ω. The members of Ω may be classed according to whether they specify $\sigma > \sigma_0$, or $\sigma < \sigma_0$. The practice of regarding only one tail of the distribution, the upper or lower depending on whether $v > N$ or $v < N$, is tantamount to accepting as admissible alternatives to H_0 only one of the classes of Ω.

The alternatives may sometimes be limited to one class or the other through some a priori knowledge, or the problem may be such that only one of the classes is relevant. However, since this is not generally the case, some method of considering all of the alternatives is needed. When testing hypotheses concerning the mean of the sampled population, the problem is quite simple, since the distribution of means is symmetrical. Thus, the "corresponding" value to any positive deviation, $(\bar{x} - m)$, is the negative deviation of the same magnitude. Merely doubling the tail area pertaining to either of the deviations will serve to take account of both classes of alternatives, i.e., those in which $m > m_0$ and those in which $m < m_0$. The problem is more difficult in the case of v or v',

[1] From the Memorial Foundation for Neuro-Endocrine Research and the Research Service of the Worcester State Hospital, Worcester, Massachusetts.

since the distribution is not symmetrical. In addition to the value of v or v' pertaining to the observed sample we require a "corresponding" value at the other end of the distribution. The definition of "corresponding" which is accepted will determine the required value. There may be a number of such definitions but not all of these will be equally acceptable. The value of v which delimits an equal tail area specifies one of the possible definitions of "corresponding." Another definition would require that the ordinates at the two values of v be equal.

The Neyman and Pearson Approach. Generalized procedures for testing statistical hypotheses have been elaborated in recent years by J. Neyman and E. S. Pearson (1–5). These have considerable philosophical appeal and will be traced as a basis of solution of the immediate problem. A test of a hypothesis H_0 consists essentially of a rule for rejecting H_0 when the observed sample E falls within a suitable critical region w of the N-dimensioned sample space W, and of accepting H_0 when E falls in $(W - w)$. In testing any hypothesis two types of error may be made:

i) H_0 may be rejected when it is true;

ii) H_0 may be accepted when some alternative hypothesis, H_i, is true.

Errors of the first kind may be considered "equivalent" since, if a true hypothesis is to be rejected, it is immaterial which one is chosen. Furthermore, the first type of error can be controlled through our choice of the size of w, say α. The size of w represents the probability of a sample E being an element of w when the hypothesis H_0 is true. This probability may be designated briefly as $P\{E \epsilon w \mid H_0\}$. Then

$$P\{E \epsilon w \mid H_0\} = \int \cdots \int_w p(E \mid H_0)\, dx_1\, dx_2 \cdots dx_N = \alpha \,.. \qquad .(\text{III})$$

where $p(E \mid H_0)$ is the elementary probability law of the sample when H_0 is true, i.e.,

$$p(E \mid H_0) = p(x_1, x_2, \cdots x_N \mid H_0) \ldots\ldots\ldots\ldots\ldots (\text{IV})$$

Errors of the second type, however, are not equivalent, since their consequences depend on the difference of the true hypothesis from H_0. The utility of a test of H_0 will depend largely on how it controls the second type of error. Ideally, the selection of a critical region should take into consideration the probabilities á priori of the hypotheses composing Ω. Since these probabilities are generally unknown, tests may be sought which are valid independently of them.

A distinction must be made between simple hypotheses which specify completely the elementary probability law of the sample, $p(E)$, and composite hypotheses which specify the law subject to one or more undetermined parameters.

(2) Simple Hypothesis Concerning Population Variance

A test based on a critical region w_0 may be called independent of the probabilities á priori of the alternative hypotheses if it is more powerful than any other

equivalent test for all of the alternative hypotheses (3). An equivalent test is one based on a region w_1 of the same size, α, i.e.,

$$P\{E \epsilon w_0 \mid H_0\} = P\{E \epsilon w_1 \mid H_0\} = \alpha \quad \text{(V)}$$

The power of a test based on any critical region, as w_1, is the probability of its rejecting a hypothesis H_0 when some other hypothesis H_i is true. That is, it is the probability of E falling in w_1 when H_i is true. Denote this power by $P\{E \epsilon w_1 \mid H_i\}$. The greater the power of a test, the smaller the risk of the second type of error. If tests as defined above exist, they minimize the probability of the second type of error. Furthermore, the probability of the first type of error is no larger than α. Neyman and Pearson (2) have designated regions satisfying this definition as Best Critical Regions for testing H_0 with regard to the set Ω. If there is no such Best Critical Region, some compromise region must be chosen.

A necessary and sufficient condition for w_0 to be a Best Critical Region with regard to an alternative H_i is that within w_0

$$p(E \mid H_0) \leq kp(E \mid H_i) \quad \text{(VI)}$$

where k is some constant depending on α. If this inequality is true for any H_i, w_0 will be a Best Critical Region for the set Ω.

Neyman and Pearson (2) have shown that in testing the hypothesis that $\sigma = \sigma_0$, when the population mean m is known, there are two Best Critical regions, one pertaining to the class of alternatives for which $\sigma < \sigma_0$ and defined by $v \leq v_1$, the other to the class $\sigma > \sigma_0$ defined by $v \geq v_2$. v_1 and v_2 are values of v so chosen that the size of the critical region shall be α. Although there is no Best Critical Region for all of the alternatives, the choice of a compromise critical region should still depend on its control of the second source of error, that is, on its power for the various alternatives (4). Such a compromise region may be designated as a Good Critical Region. What is needed is a region w_0 of size α defined by the inequalities $v \leq v_1$ and $v \geq v_2$. If v_1 and v_2 are taken as the values cutting off equal tail areas, then the power of the test will be less than α for some values of σ less than σ_0. For those values of σ, H_0 would be accepted more frequently than if it were true. Thus a first requirement for a Good Critical Region is that its power should nowhere be less than α, the value when H_0 is true. Of all such unbiassed Critical Regions of size α, w_0 should then be selected so that its power is everywhere greater than that of any other equivalent unbiassed region.

Critical Regions sufficiently satisfying the above requirements can often be obtained by stipulating that the first derivative of the power function with respect to θ, the parameter under consideration, shall be zero at $\theta = \theta_0$, and that the second shall be a maximum there. Then not only does the probability of the second source of error decrease as we move away from θ_0, but it decreases most rapidly in the vicinity of θ_0. Critical Regions satisfying these conditions are called unbiassed Critical Regions of Type A, (4). Under certain assumptions

concerning the nature of the elementary probability law $p(E \mid \theta)$ it can be shown that w_0 is defined by the inequalities $\varphi_1 \leq c_1$ and $\varphi_1 \geq c_2$ where c_1 and c_2 satisfy the conditions

$$\int_{c_1}^{c_2} p(\varphi_1)\, d\varphi_1 = 1 - \alpha \qquad \text{(VII)}$$

$$\int_{c_1}^{c_2} \varphi_1 p(\varphi_1)\, d\varphi_1 = 0 \qquad \text{(VIII)}$$

where
$$\varphi_1 = \frac{d \log p(E \mid \theta)}{d\theta}\bigg|_{\theta=\theta_0} \qquad \text{(IX)}$$

and $p(\varphi_1)$ is the distribution function of φ_1.

In applying these results to the testing of the hypothesis that $\sigma = \sigma_0$[2] when the population mean is known,

$$\varphi_1 = (v - N)/\sigma_0 \qquad \text{(X)}$$

Obviously $p(v)$, the distribution of v, may be considered instead of $p(\varphi_1)$. w_0 is defined by the inequalities $v \leq v_1$ and $v \geq v_2$ where

$$\int_0^{v_1} p(v)\, dv + \int_{v_2}^{\infty} p(v)\, dv = \alpha_1 + \alpha_2 = \alpha \qquad \text{(XI)}$$

$$\int_{v_1}^{v_2} (v - N)p(v)\, dv = v^{N/2}e^{-v/2}\Big|_{v_1}^{v_2} = 0 \qquad \text{(XII)}$$

w_0 so defined is also of type A_1, that is, its power curve lies everywhere above that of any other equivalent region, vanishing in the first derivative at $\sigma = \sigma_0$, (4).

The use of w_0 as the appropriate critical region is equivalent to the use of r as a test criterion, where

$$v^{N/2}e^{-\frac{1}{2}v} = r) \qquad \text{XIII)}$$

That is, a value of v yielding the same r as the observed v may be taken as the corresponding value. Reference to the appropriate tables and summing of the two tail areas gives P_r, the probability of obtaining a smaller value of r when H_0 is true. H_0 may be rejected if P_r is less than some previously fixed number, say α. If the distribution of r could be evaluated the necessity of dealing with two values of v would be obviated.

The criterion r is equivalent to that deduced by the use of maximum likelihood ratios (6). Thus,

$$p(E \mid \sigma^2) = (2\pi\sigma^2)^{-N/2} e^{-\overset{N}{\underset{i=1}{S}} (x_i - m)^2/2\sigma^2} \qquad \text{(XIV)}$$

[2] The solution is the same in terms of σ^2.

Maximizing $p(E \mid \sigma^2)$ for fixed E and all possible σ^2 we have

$$p_{\max.}(E \mid \sigma^2) = \ldots \left[\ldots \sum_{i=1} \ldots \right] \ldots$$

$$\lambda = \frac{p(E \mid \sigma_0^2)}{p_{\max.}(E \mid \sigma^2)} = N^{-N/2} v^{N/2} e^{-\frac{1}{2}(v-N)} \qquad \text{(XVI)}$$

$$= N^{-N/2} e^{N/2} r \qquad \text{(XVII)}$$

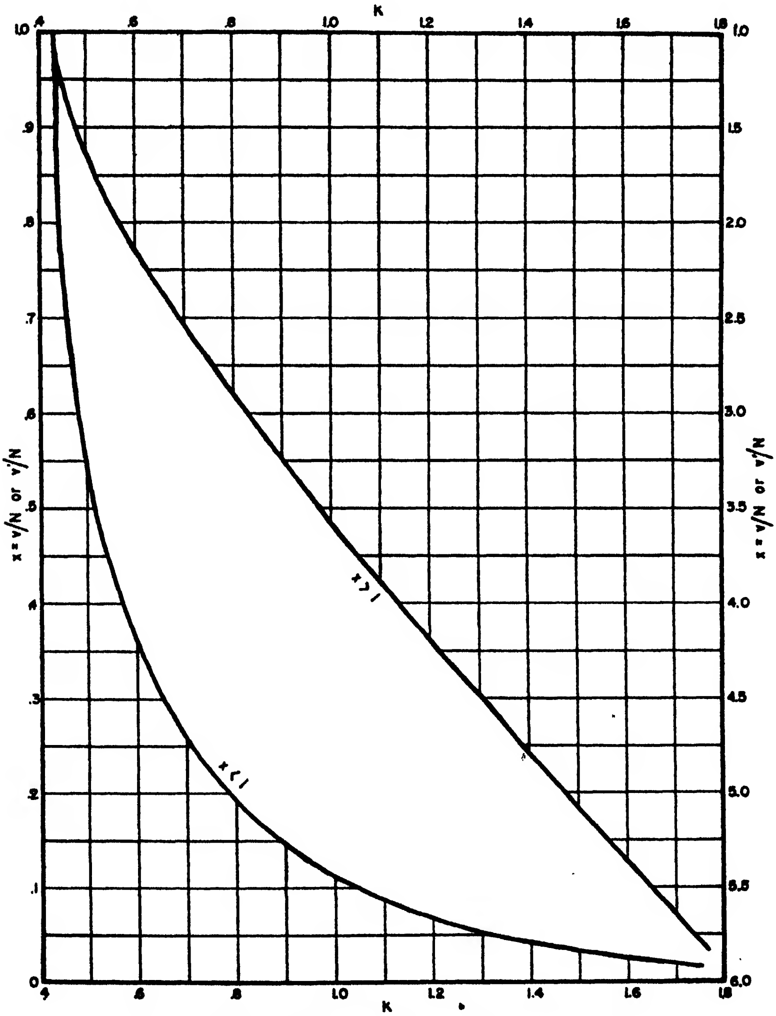


FIG. 1. Graph of Equation $x - \log_e x = k \log_e 10$

The h^{th} moment coefficient of λ about zero, $\mu_h'(\lambda)$, is given by

$$\mu_h'(\lambda) = \frac{\Gamma\left[\dfrac{N(1+h)}{2}\right]}{\Gamma(N/2)} (2e/N)^{hN/2} (1+h)^{-N(1+h)/2} \qquad \text{(X}$$

TABLE I

Probability that a sample has been drawn from a normal population with a specified variance or standard deviation

Degrees of Freedom, n

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0.435	.9724	.9581	.9473	.9383	.9305	.9235	.9171	.9111	.9055	.9002	.8952	.8905	.8859	.8815	.8773	.8732	.8693	.8655	.8617	.8581	.8546	.8512	.8478	.8445	.8413
0.440	.9217	.8812	.8509	.8259	.8042	.7848	.7671	.7508	.7357	.7215	.7081	.6954	.6833	.6717	.6606	.6500	.6398	.6299	.6204	.6112	.6023	.5936	.5853	.5771	.5692
0.445	.8928	.8377	.7968	.7632	.7341	.7083	.6850	.6636	.6438	.6253	.6080	.5917	.5762	.5616	.5476	.5343	.5215	.5093	.4975	.4862	.4753	.4649	.4547	.4450	.4355
0.450	.8704	.8041	.7552	.7151	.6808	.6505	.6232	.5983	.5754	.5542	.5344	.5159	.4984	.4820	.4664	.4516	.4375	.4241	.4113	.3990	.3873	.3760	.3652	.3549	.3449
0.455	.8513	.7758	.7203	.6752	.6367	.6029	.5726	.5452	.5202	.4971	.4756	.4557	.4370	.4195	.4030	.3874	.3726	.3587	.3454	.3328	.3208	.3094	.2985	.2881	.2781
0.460	.8346	.7510	.6899	.6405	.5987	.5621	.5296	.5003	.4737	.4492	.4267	.4059	.3865	.3683	.3514	.3355	.3205	.3064	.2931	.2806	.2687	.2574	.2467	.2366	.2269
0.465	.8194	.7287	.6628	.6098	.5651	.5263	.4920	.4613	.4335	.4082	.3850	.3636	.3439	.3255	.3084	.2925	.2776	.2637	.2506	.2383	.2267	.2158	.2055	.1958	.1866
0.470	.8055	.7083	.6382	.5821	.5350	.4944	.4587	.4270	.3984	.3725	.3489	.3273	.3074	.2890	.2721	.2563	.2417	.2281	.2154	.2035	.1924	.1819	.1722	.1630	.1544
0.475	.7926	.6896	.6156	.5568	.5077	.4657	.4289	.3963	.3672	.3410	.3172	.2956	.2758	.2576	.2409	.2255	.2113	.1981	.1859	.1745	.1639	.1541	.1449	.1364	.1284
0.480	.7805	.6721	.5946	.5335	.4827	.4395	.4019	.3688	.3393	.3130	.2892	.2677	.2481	.2303	.2140	.1990	.1853	.1726	.1610	.1502	.1402	.1310	.1224	.1145	.1071
0.485	.7692	.6557	.5751	.5119	.4597	.4155	.3773	.3438	.3142	.2879	.2643	.2430	.2238	.2064	.1906	.1761	.1630	.1509	.1398	.1296	.1203	.1117	.1037	.0964	.0897
0.490	.7583	.6402	.5569	.4918	.4384	.3934	.3547	.3211	.2915	.2653	.2420	.2211	.2023	.1854	.1701	.1562	.1436	.1322	.1217	.1122	.1035	.0955	.0881	.0814	.0753
0.495	.7481	.6256	.5397	.4729	.4185	.3729	.3340	.3003	.2708	.2449	.2219	.2015	.1832	.1668	.1521	.1388	.1269	.1160	.1062	.0973	.0892	.0818	.0750	.0689	.0633
0.500	.7382	.6117	.5234	.4552	.4000	.3539	.3148	.2812	.2519	.2263	.2038	.1838	.1661	.1503	.1362	.1236	.1122	.1020	.0928	.0845	.0770	.0702	.0640	.0584	.0534
0.510	.7197	.5857	.4933	.4228	.3663	.3197	.2806	.2473	.2188	.1940	.1725	.1537	.1372	.1226	.1097	.0983	.0882	.0792	.0712	.0640	.0577	.0519	.0468	.0422	.0381
0.520	.7025	.5619	.4660	.3937	.3364	.2897	.2509	.2183	.1907	.1670	.1466	.1290	.1138	.1004	.0888	.0786	.0697	.0618	.0549	.0488	.0434	.0386	.0344	.0307	.0273
0.530	.6864	.5399	.4411	.3674	.3097	.2632	.2250	.1933	.1667	.1442	.1251	.1087	.0947	.0826	.0721	.0631	.0553	.0484	.0425	.0373	.0328	.0289	.0254	.0224	.0197
0.540	.6713	.5194	.4181	.3435	.2856	.2396	.2023	.1716	.1461	.1248	.1070	.0918	.0790	.0681	.0588	.0508	.0439	.0381	.0330	.0286	.0249	.0216	.0188	.0164	.0143
0.550	.6570	.5002	.3969	.3216	.2639	.2186	.1822	.1526	.1284	.1083	.0917	.0778	.0661	.0563	.0480	.0410	.0351	.0300	.0257	.0221	.0189	.0163	.0140	.0120	.0103
0.560	.6434	.4822	.3772	.3015	.2442	.1997	.1643	.1360	.1130	.0942	.0787	.0660	.0554	.0466	.0393	.0332	.0280	.0237	.0201	.0170	.0144	.0123	.0104	.0089	.0075
0.570	.6304	.4652	.3588	.2830	.2263	.1827	.1485	.1213	.0996	.0820	.0678	.0561	.0466	.0387	.0322	.0269	.0224	.0188	.0157	.0132	.0110	.0093	.0078	.0065	.0055
0.580	.6180	.4492	.3417	.2659	.2099	.1673	.1343	.1084	.0879	.0715	.0584	.0478	.0392	.0322	.0265	.0218	.0180	.0149	.0123	.0102	.0084	.0070	.0058	.0048	.0040
0.590	.6061	.4340	.3256	.2501	.1949	.1534	.1217	.0970	.0777	.0625	.0504	.0407	.0330	.0268	.0218	.0177	.0145	.0118	.0097	.0079	.0065	.0053	.0044	.0036	.0029
0.600	.5946	.4195	.3105	.2354	.1811	.1408	.1103	.0869	.0688	.0546	.0435	.0348	.0278	.0223	.0180	.0145	.0117	.0094	.0076	.0061	.0050	.0040	.0033	.0027	.0022
0.610	.5836	.4057	.2963	.2217	.1685	.1294	.1001	.0779	.0609	.0478	.0376	.0297	.0235	.0186	.0148	.0118	.0094	.0075	.0060	.0048	.0038	.0031	.0025	.0020	.0016
0.620	.5730	.3926	.2829	.2090	.1568	.1190	.0909	.0699	.0540	.0419	.0326	.0254	.0199	.0156	.0122	.0096	.0076	.0060	.0047	.0037	.0029	.0023	.0018	.0015	.0012
0.630	.5627	.3801	.2702	.1971	.1461	.1094	.0826	.0628	.0479	.0367	.0282	.0218	.0168	.0130	.0101	.0079	.0061	.0048	.0037	.0029	.0023	.0018	.0014	.0011	.0009
0.640	.5528	.3682	.2583	.1860	.1362	.1008	.0752	.0564	.0426	.0322	.0245	.0187	.0143	.0109	.0084	.0064	.0049	.0038	.0029	.0023	.0018	.0014	.0010	.0008	.0006

0.650	.5432	.3567	.2470	.1757	.1270	.0928	.0684	.0508	.0378	.0283	.0213	.0160	.0121	.0091	.0069	.0053	.0040	.0030	.0023	.0018	.0014	.0010	.0008	.0006	.0005
0.660	.5339	.3457	.2363	.1659	.1185	.0856	.0623	.0457	.0336	.0249	.0185	.0137	.0103	.0077	.0057	.0043	.0032	.0024	.0018	.0014	.0010	.0008	.0006	.0005	.0003
0.670	.5249	.3352	.2261	.1568	.1106	.0789	.0568	.0411	.0299	.0219	.0160	.0118	.0087	.0064	.0048	.0035	.0026	.0020	.0015	.0011	.0008	.0006	.0005	.0003	.0003
0.680	.5161	.3251	.2165	.1483	.1033	.0728	.0518	.0371	.0267	.0193	.0140	.0101	.0074	.0054	.0040	.0029	.0021	.0016	.0012	.0008	.0006	.0005	.0003	.0003	.0002
0.690	.5076	.3154	.2073	.1403	.0965	.0672	.0472	.0334	.0237	.0169	.0121	.0087	.0063	.0045	.0033	.0024	.0017	.0013	.0009	.0007	.0005	.0004	.0003	.0002	.0001
0.700	.4993	.3060	.1986	.1327	.0902	.0621	.0431	.0301	.0212	.0149	.0106	.0075	.0053	.0038	.0027	.0020	.0014	.0010	.0007	.0005	.0004	.0003	.0002	.0001	.0001
0.750	.4609	.2642	.1610	.1011	.0647	.0419	.0274	.0181	.0120	.0080	.0053	.0036	.0024	.0016	.0011	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000	.0000	.0000
0.800	.4268	.2292	.1312	.0776	.0468	.0286	.0176	.0109	.0068	.0043	.0027	.0017	.0011	.0007	.0004	.0003	.0002	.0001	.0001	.0000	.0000	.0000			
0.850	.3962	.1995	.1074	.0598	.0339	.0195	.0114	.0067	.0039	.0023	.0014	.0008	.0005	.0003	.0002	.0001	.0001	.0000	.0000						
0.900	.3686	.1742	.0883	.0463	.0248	.0134	.0074	.0041	.0023	.0013	.0007	.0004	.0002	.0001	.0001	.0000	.0000								
0.950	.3435	.1525	.0727	.0359	.0181	.0093	.0048	.0025	.0013	.0007	.0004	.0002	.0001	.0001	.0000										
1.000	.3205	.1338	.0601	.0280	.0133	.0064	.0031	.0015	.0008	.0004	.0002	.0001	.0000	.0000											
1.050	.2994	.1175	.0497	.0218	.0098	.0045	.0020	.0010	.0004	.0002	.0001	.0000													
1.100	.2800	.1034	.0412	.0170	.0072	.0031	.0013	.0006	.0003	.0001	.0001														
1.150	.2621	.0911	.0342	.0133	.0053	.0022	.0009	.0004	.0002	.0001	.0000														
1.200	.2455	.0803	.0284	.0105	.0039	.0015	.0006	.0002	.0001	.0000															
1.250	.2301	.0709	.0236	.0082	.0029	.0011	.0004	.0001	.0001																
1.300	.2158	.0626	.0197	.0064	.0022	.0007	.0003	.0001	.0000																
1.350	.2024	.0553	.0164	.0051	.0016	.0005	.0002	.0001																	
1.400	.1900	.0490	.0137	.0040	.0012	.0004	.0001	.0000																	
1.450	.1785	.0433	.0114	.0031	.0009	.0003	.0001																		
1.500	.1677	.0384	.0096	.0025	.0007	.0002	.0000																		
1.800	.1159	.0187	.0033	.0006	.0001	.0000																			
2.100	.0807	.0092	.0011	.0001	.0000																				
2.400	.0564	.0045	.0004	.0000																					
2.700	.0395	.0022	.0001																						
3.000	.0278	.0011	.0000																						

TABLE I—*Concluded*

k	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
0.435	.8382	.8351	.8321	.8291	.8262	.8234	.8206	.8178	.8151	.8124	.8098	.8072	.8046	.8021	.7996	.7972	.7947	.7923	.7900	.7876	.7853	.7831	.7808	.7786	.7764
0.440	.5614	.5539	.5466	.5394	.5324	.5256	.5189	.5124	.5060	.4998	.4936	.4876	.4818	.4760	.4704	.4648	.4594	.4540	.4488	.4437	.4386	.4336	.4287	.4239	.4192
0.445	.4263	.4175	.4089	.4005	.3924	.3846	.3769	.3695	.3623	.3553	.3484	.3417	.3352	.3289	.3227	.3167	.3108	.3051	.2995	.2940	.2887	.2835	.2784	.2734	.2685
0.450	.3353	.3261	.3172	.3086	.3003	.2923	.2845	.2771	.2698	.2628	.2561	.2495	.2432	.2370	.2310	.2252	.2196	.2142	.2089	.2037	.1988	.1939	.1892	.1846	.1802
0.455	.2686	.2595	.2507	.2424	.2343	.2266	.2192	.2120	.2052	.1986	.1922	.1861	.1802	.1745	.1691	.1638	.1587	.1538	.1490	.1445	.1401	.1358	.1317	.1277	.1238
0.460	.2177	.2090	.2006	.1927	.1851	.1779	.1710	.1644	.1580	.1520	.1462	.1407	.1354	.1303	.1254	.1208	.1163	.1120	.1079	.1039	.1002	.0965	.0930	.0896	.0864
0.465	.1779	.1697	.1619	.1545	.1475	.1409	.1346	.1286	.1229	.1174	.1123	.1074	.1027	.0982	.0940	.0899	.0861	.0824	.0789	.0755	.0723	.0693	.0664	.0636	.0609
0.470	.1463	.1387	.1315	.1247	.1183	.1123	.1066	.1012	.0962	.0914	.0868	.0825	.0784	.0746	.0709	.0675	.0642	.0611	.0581	.0553	.0527	.0501	.0477	.0455	.0433
0.475	.1209	.1139	.1073	.1012	.0954	.0900	.0849	.0801	.0757	.0714	.0675	.0638	.0602	.0569	.0538	.0509	.0481	.0455	.0431	.0407	.0386	.0365	.0345	.0327	.0310
0.480	.1002	.0939	.0879	.0824	.0772	.0724	.0679	.0637	.0598	.0561	.0527	.0495	.0465	.0437	.0410	.0386	.0362	.0341	.0320	.0301	.0284	.0267	.0251	.0236	.0222
0.485	.0834	.0776	.0723	.0673	.0627	.0584	.0545	.0508	.0474	.0442	.0413	.0385	.0360	.0336	.0314	.0293	.0274	.0256	.0239	.0224	.0209	.0196	.0183	.0171	.0160
0.490	.0696	.0644	.0596	.0551	.0511	.0473	.0439	.0406	.0377	.0350	.0324	.0301	.0279	.0259	.0241	.0224	.0208	.0193	.0179	.0167	.0155	.0144	.0134	.0125	.0116
0.495	.0582	.0535	.0492	.0453	.0417	.0384	.0354	.0326	.0300	.0277	.0256	.0236	.0218	.0201	.0185	.0171	.0158	.0146	.0135	.0125	.0115	.0106	.0098	.0091	.0084
0.500	.0487	.0446	.0407	.0373	.0341	.0312	.0286	.0262	.0240	.0220	.0202	.0185	.0170	.0156	.0143	.0131	.0120	.0111	.0102	.0093	.0086	.0079	.0072	.0067	.0061
0.510	.0344	.0311	.0281	.0254	.0229	.0208	.0188	.0170	.0154	.0140	.0126	.0115	.0104	.0094	.0085	.0078	.0070	.0064	.0058	.0053	.0048	.0043	.0039	.0036	.0033
0.520	.0244	.0218	.0194	.0174	.0155	.0139	.0124	.0111	.0099	.0089	.0080	.0071	.0064	.0057	.0051	.0046	.0041	.0037	.0033	.0030	.0027	.0024	.0022	.0019	.0017
0.530	.0174	.0153	.0135	.0119	.0106	.0093	.0082	.0073	.0064	.0057	.0051	.0045	.0040	.0035	.0031	.0028	.0024	.0022	.0019	.0017	.0015	.0013	.0012	.0011	.0009
0.540	.0124	.0108	.0095	.0082	.0072	.0063	.0055	.0048	.0042	.0037	.0032	.0028	.0025	.0022	.0019	.0017	.0014	.0013	.0011	.0010	.0009	.0008	.0007	.0006	.0005
0.550	.0089	.0077	.0066	.0057	.0049	.0043	.0037	.0032	.0027	.0024	.0021	.0018	.0015	.0013	.0011	.0010	.0009	.0007	.0006	.0006	.0005	.0004	.0004	.0003	.0003
0.560	.0064	.0055	.0047	.0040	.0034	.0029	.0025	.0021	.0018	.0015	.0013	.0011	.0010	.0008	.0007	.0006	.0005	.0004	.0004	.0003	.0003	.0002	.0002	.0002	.0001
0.570	.0046	.0039	.0033	.0028	.0023	.0020	.0017	.0014	.0012	.0010	.0008	.0007	.0006	.0005	.0004	.0004	.0003	.0003	.0002	.0002	.0002	.0001	.0001	.0001	.0001
0.580	.0033	.0028	.0023	.0019	.0016	.0013	.0011	.0009	.0008	.0006	.0005	.0005	.0004	.0003	.0003	.0002	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0000
0.590	.0024	.0020	.0016	.0013	.0011	.0009	.0008	.0006	.0005	.0004	.0003	.0003	.0002	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000	
0.600	.0017	.0014	.0012	.0009	.0008	.0006	.0005	.0004	.0003	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000				
0.610	.0013	.0010	.0008	.0007	.0005	.0004	.0003	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0000	.0000	.0000							
0.620	.0009	.0007	.0006	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0000	.0000										
0.630	.0007	.0005	.0004	.0003	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000												
0.640	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0000														
0.650	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000															
0.660	.0003	.0002	.0001	.0001	.0001	.0001	.0000	.0000																	
0.670	.0002	.0001	.0001	.0001	.0001	.0000																			
0.680	.0001	.0001	.0001	.0001	.0000																				
0.690	.0001	.0001	.0001	.0000																					
0.700	.0001	.0001	.0000																						
0.750	.0000	.0000																							

For N infinite, $(-2\log_e\lambda)$ will be distributed as χ^2 with one degree of freedom. For finite values of N, however, we have not been able to evaluate the distribution of λ, although the distribution of the Incomplete Beta Function serves as a good approximation. Approximate distributions for several values of N have been obtained. P_λ, the probability of obtaining a smaller value of λ than that observed, as obtained from these distributions agrees well with the sum of the tail areas pertaining to v_1 and v_2 yielding the same value of λ (or r). The construction of tables is simplified by taking (1)

$$\log_{10} \lambda = N/2(\log_{10} e - k) \quad \text{(XIX)}$$

That is,

$$x - \log_e x = k \log_e 10 \quad \text{(XX)}$$

where $x = v/N$. Equation (XX) is independent of N and may be solved once and for all for x, given k.[3] In Figure 1 is plotted the graph of equation (XX). For convenience, the branch of the curve giving the roots greater than unity has been folded back with altered scale from the minimum value of k, $\log_{10}e$, occurring at $x = 1$. Table I was then constructed by multiplying the two values of x for a given k by $(N/2)^{\frac{1}{2}}$, referring to the Tables of the Incomplete Gamma Function (7) with $p = (N - 2)/2$, and adding the resulting two tail areas. The values for the odd numbers above 12 were obtained by interpolating between the even numbers. For $N = 1$, $(x)^{\frac{1}{2}}$ was used as a normal deviate. The values in Table I should be correct to four decimals. Table I is entered with the number of degrees of freedom, n, on which x is based. In the case of the simple hypothesis this is N.

The following may serve as an illustration: Blood urea nitrogen determinations (mg./100 cc.) were made on a sample of 25 schizophrenic patients. The mean was found to be 15.56, the variance, 10.486. Previous investigation of blood urea nitrogen on a large sample of normal control subjects gave a mean of 16.03 and a variance of 20.268, which for the purpose of the example may be considered as the population parameters. Then we may wish to test the hypothesis that the variance of the sampled population, σ^2, is $\sigma_0^2 = 20.268$, knowing the mean of the sampled population to be 16.03. Calculate

$$\frac{s^2 + (\bar{x} - m)^2}{\sigma_0^2} \quad .528$$

Referring to Fig. 1, the value of k is about .505. Turning to Table I with $k = .505$,[4] $n = 25$, P is found to be .0457. We should thus be inclined to reject the hypothesis.

For N small, the area of the tail of the distribution near zero is considerably larger than that at the upper end. As N increases the distribution of v becomes

---

[3] If the solution were explicit the distribution of λ could easily be deduced from that of x.

[4] k obtained directly from (XX) is .507, corresponding to $P = .0427$.

more and more symmetrical and the two areas approach equality. Even for $N = 50$, however, they are rather unequal, so that merely doubling the area pertaining to the observed v does not give a sufficiently accurate approximation. For $N > 50$ an approximation correct within several units in the third decimal place may be obtained by taking $\sqrt{2N}(\sqrt{x} - 1)$ as a normal deviate. This assumes that the standard deviation is normally distributed with variance $\sigma_0^2/2N$.

(3) Composite Hypothesis Concerning Population Variance

Here H_0 specifies only the value of the parameter $\theta = \theta_0$, leaving undetermined the value of a second parameter, ν. Thus, H_0 consists of a subset, ω, of simple hypotheses, each of which specifies a different value for ν. Any simple hypothesis specifying different values of both parameters, θ and ν, is an alternative to H_0. These alternatives form the set Ω. The elementary probability law determined by H_0 is $p(E \mid H_0) = p(E \mid \theta_0\nu)$, while that determined by an alternative hypothesis H_i is $p(E \mid H_i) = p(E \mid \theta_i\nu_i)$. In testing composite hypotheses the first requirement is to find regions "similar" to W with regard to ν, i.e., such that the chance of rejection of a true hypothesis, $P\{E \epsilon w \mid H_0\}$, equals α for all the values of ν specified by the simple hypotheses composing H_0. A test based on a similar region w_0 may be called independent of the probabilities á priori, if its power with respect to all the alternatives of Ω is greater than that of any other similar region w_1 of the same size, α, (3). Let

$$\varphi_2 = \partial \log p(E \mid \theta\nu)/\partial\nu \big|_{\theta=\theta_0} \ldots\ldots\ldots\ldots\ldots\ldots (\mathrm{XXI})$$

Then the equations $\varphi_2 =$ constant will describe hypersurfaces in N-dimensioned space, on one of which the observed E must fall. Under certain assumptions pertaining to the law of elementary probability it can be shown (2) that a necessary and sufficient condition for w to be a similar region is that

$$P\{E \epsilon w(\varphi_2) \mid H_0\} = \alpha P\{E \epsilon W(\varphi_2) \mid H_0\} \ldots\ldots\ldots\ldots (\mathrm{XXII})$$

for all values of φ_2, where $w(\varphi_2)$ and $W(\varphi_2)$ are parts of the surface $\varphi_2 =$ constant common to w and W respectively. A similar region is then built up of these parts $w(\varphi_2)$ obtaining for the various values of φ_2. The Best Critical Region, w_0, for a particular simple alternative, H_i, must then be composed of pieces, $w_0(\varphi_2)$, maximizing $P\{E \epsilon w_0(\varphi_2) \mid H_i\}$. The problem is the same as for simple hypotheses except that we shall be working in a space $W(\varphi_2)$ of $(N - 1)$ dimensions. $w_0(\varphi_2)$ is defined by the inequality

$$p(E \mid H_i) \geq k(\varphi_2)\; p(E \mid H_0) \ldots\ldots\ldots\ldots\ldots (\mathrm{XXIII})$$

where $k(\varphi_2)$ is some constant depending on α. If $w_0(\varphi_2)$ is the same for all H_i, then w_0 is the Best Critical Region for testing H_0 with respect to Ω.

Neyman and Pearson showed (2) that in testing the composite hypothesis that $\sigma = \sigma_0$ when the population mean is unknown there are two Best Critical Regions corresponding to the class of alternatives $\sigma < \sigma_0$ and $\sigma > \sigma_0$, defined respectively by the inequalities $v' \leq v_1'$ and $v' \geq v_2'$. If the whole set of alternatives, Ω, is to

be considered some compromise region must be sought. Dealing with the case where similar regions exist Neyman (5) defines a Critical Region as unbiassed and of Type B if the first derivative of the power function, $P(E \epsilon w \mid H_i)$, with respect to θ vanishes at $\theta = \theta_0$, and if the second derivative at that point is a maximum. Let

$$\varphi_1 = \frac{\partial \log p(E \mid \theta\nu)}{\partial \theta}\bigg|_{\theta=\theta_0} \quad \text{.(XXIV)}$$

Then it can be shown that the desired region will be defined by the inequalities $\varphi_1 \leq k_1(\varphi_2)$ and $\varphi_1 \geq k_2(\varphi_2)$ where $k_1(\varphi_2)$ and $k_2(\varphi_2)$ are determined to satisfy

$$\int_{k_1(\varphi_2)}^{k_2(\varphi_2)} p(\varphi_1 \varphi_2)\, d\varphi_1 = (1 - \alpha)p(\varphi_2) \quad \text{.............(XXV)}$$

and

$$\int_{k_1(\varphi_2)}^{k_2(\varphi_2)} \varphi_1 p(\varphi_1 \varphi_2)\, d\varphi_1 = (1 - \alpha) \int_{-\infty}^{\infty} \varphi_1 p(\varphi_1 \varphi_2)\, d\varphi_1 \quad \text{......(XXVI)}$$

where $p(\varphi_2)$ is the distribution function of φ_2, and $p(\varphi_1\varphi_2)$ is the simultaneous distribution of φ_1 and φ_2.

Applying equations (XXV) and (XXVI) it follows that the appropriate Critical Region is defined by the inequalities $v' \leq v_1'$ and $v' \geq v_2'$ where

$$\alpha = \alpha_1 + \alpha_2 = \int_0^{v_1'} p(v')\, dv' + \int_{v_2'}^{\infty} p(v')\, dv' \quad \text{.......(XXVII)}$$

and

$$v'^{(N-1)/2} e^{-\frac{1}{2}v'} \Big|_{v_1'}^{v_2'} = 0 \quad \text{.................(XXVIII)}$$

where $p(v')$ is the distribution function of v'.

The use of the unbiassed Critical Region of Type B corresponds to adopting as a criterion

$$v'^{(N-1)/2} e^{-\frac{1}{2}v'} = r' \quad \text{..................(XXIX)}$$

Since v' derived from a sample of size N is distributed as v derived from a sample of size $(N - 1)$, it follows that r' is equivalent to the r of equation (XIII) based on a sample of size $(N - 1)$. Therefore Table I may also be used for testing the hypothesis that $\sigma = \sigma_0$ whatever be the population mean, by entering with the number of degrees of freedom, $N - 1$.

In the example previously used, compute

$$x = \frac{s^2}{\sigma_0^2} = 0.517$$

From Figure 1, k is approximately .51, corresponding to $P = .0422$.

r' is not the same as the maximum likelihood ratio λ' (6).

$$\lambda' = \frac{p_{\max}(E \mid \sigma_0^2 m)}{p_{\max}(E \mid \sigma^2 m)} = N^{-N/2} v'^{N/2} e^{-\frac{1}{2}(v'-N)} = N^{-N/2} e^{N/2} v'^{\frac{1}{2}} r' \ldots \text{(XXX)}$$

As N becomes infinite the distribution of λ' is the same as that of the λ of (XVI). For $N = 49$, the probabilities corresponding to λ' agree with those using r' to within a unit in the third decimal.

The λ' test is biassed as may be seen in Figure 2 where we have plotted the power of the test based on the region w defined by $v_1' = 3.187$, $v_2' = 22.912$ for which $\alpha = .0436 + .0064 = .0500$, on the assumption that $\sigma_0^2 = 1.0$, for $N = 10$. Although the criterion is biassed it is slightly more sensitive to alternatives

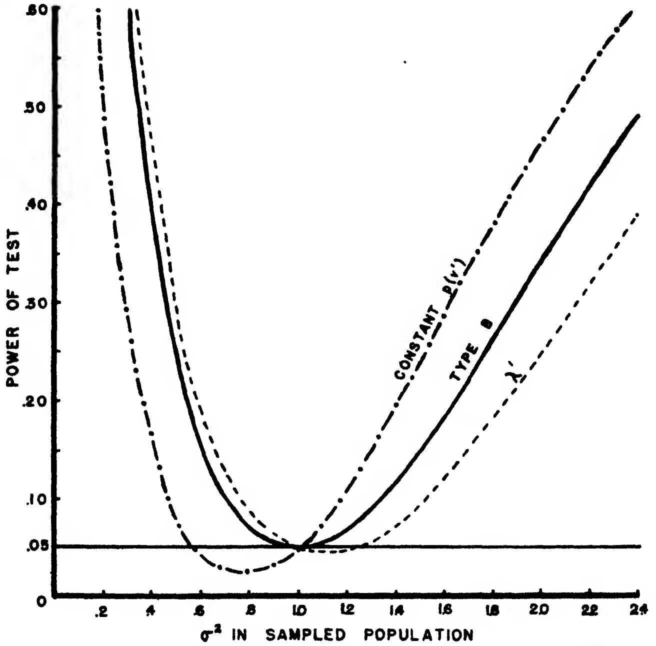


FIG. 2. Comparison of Critical Regions for v'. H_0 Specifies $\sigma_0^2 = 1.0$. $N = 10$.

specifying $\sigma^2 < \sigma_0^2$ than is the unbiassed Critical Region of Type B defined by $v_1' = 2.953$, $v_2' = 20.305$, $\alpha = .0339 + .0161 = .0500$. The criterion of constant distribution, $p(v')$,

$$v'^{(N-3)/2} e^{-\frac{1}{2}v'} = c' \ldots\ldots\ldots\ldots\ldots\ldots \text{(XXXI)}$$

has also been considered. In this case $v_1' = 1.903$, $v_2' = 17.391$, $\alpha = .0071 + .0429 = .0500$. This criterion is biassed for some alternatives specifying $\sigma^2 < \sigma_0^2$, but its power curve lies above that of the unbiassed region for $\sigma^2 > \sigma_0^2$.

Apparently the bias may be shifted at will by changing the exponent of v'. This may be desirable if greater weight is to be given to one class of alternatives. In fact decreasing the exponent of v' to 0 produces the Best Critical Region

for the class of alternatives specifying $\sigma^2 > \sigma_0^2$, and defined by $v_1 = 0$, $v_2 = 16.919$ for $\alpha = .0500$. No region can be found giving greater power. On the other hand this region is insensitive to alternatives of the other class. Increasing the exponent indefinitely produces the Best Critical Region for the other class defined by $v_2' = \infty$ and $v_1' = 3.325$ for $\alpha = .0500$.

REFERENCES

(1) NEYMAN, J. AND PEARSON, E. S. Biometrika, vol. 20, 1928, pp. 175–241.
(2) Ibid. Phil. Trans. Roy. Soc. A, vol. 231, 1933, pp. 289–337.
(3) Ibid. Proc. Camb. Phil. Soc., vol. 29, 1933, pp. 492–510.
(4) Ibid. Stat. Research Memoirs, vol. 1, 1936, pp. 1–37.
(5) NEYMAN, J. Bull. d. l. Soc. Math. d. France, vol. 63, 1935, pp. 246–266.
(6) FERTIG, J. Biom. Bull., vol. 1, 1936, pp. 45–82.
(7) Tables of the Incomplete Γ-Function, 1934.

ON THE POLYNOMIALS RELATED TO THE DIFFERENTIAL EQUATION $\frac{1}{y}\frac{dy}{dx} = \frac{a_0 + a_1 x}{b_0 + b_1 x + b_2 x^2} \equiv \frac{N}{D}$

By Frank S. Beale

Introduction. In a previous issue of this Journal,[1] E. H. Hildebrandt has established the existence of a general system of polynomials $P_n(k, x)$ associated with the solutions of Pearson's Differential Equation

$$\text{(R)} \qquad \frac{1}{y}\frac{dy}{dx} = \frac{N}{D},$$

N and D being polynomials in x of degrees not exceeding one and two respectively with no factor in common.

It was shown that the polynomials $P_n(k, x) \equiv P_n$ themselves satisfy certain differential equations and a recurrence relation. The classical polynomials of Hermite, Legendre, Laguerre, and Jacobi are special types of $P_n(k, x)$. Since the classical polynomials are employed rather extensively in statistical theory, certain of their properties are of special interest.

It is the purpose of this paper to determine from Hildebrandt's general equations some new properties of $P_n(k, x)$ and to apply these properties to the classical polynomials. The paper consists of two parts. In part I some theorems are established concerning common zeros of D and P_n. In particular, a theorem is established to exhibit the conditions under which the zeros of P_n, which are not zeros of D, are simple. In part II a method is outlined for the classical polynomials by which one can determine the number and location of the real zeros in the various segments into which the zeros of D divide the x axis. The points of inflexion and the degree of the polynomials are also considered.

A new feature of the method employed is, we believe, its being based upon the use of differential equations of first order, for most part, while other investigators[2] have employed differential equations of second order. As to the results obtained, the author believes them to be partly new. They have points in common with the results of Fujiwara, Lawton and Webster.

[1] Systems of Polynomials Connected with the Charlier Expansions, etc., Annals of Math. Stat., Vol. II, 1931, pp. 379–439.

[2] *M. Fujiwara:* On the zeros of Jacobi's Polynomials, Japanese Journal of Math., Vol. 2, 1925, pp. 1, 2.

W. Lawton: On the zeros of Certain Polynomials Related to Jacobi and Laguerre Polynomials, Bull. Am. Math. Soc., Vol. 38, 1932, pp. 442–449.

M. S. Webster: Thesis, Univ. of Penna. These results were kindly communicated to me by Dr. Webster.

I. Theorems Concerning Common Zeros of $P_n(k, x)$ and D

The following equations will be employed later:

$$P_{n+1}(k, x) = [N + (k - n)D']P_n(k, x) + DP'_n(k, x). \tag{1}$$

$$P'_{n+1}(k, x) = (n + 1)\left[N' + \frac{2k - n}{2} D''\right]P_n(k, x). \tag{2}$$

$$\begin{aligned} P_{n+1}(k, x) &= [N + (k - n)D']P_n(k, x) \\ &\quad + n\left[N' + \frac{2k - n + 1}{2} D''\right]DP_{n-1}(k, x). \end{aligned} \tag{3}$$

These are not explicitly given in Hildebrandt's Paper but the method of obtaining them is outlined there in detail.

We shall make use of the following lemma which we state without proof.

Lemma (1). *Let $P_n(x)$ be a polynomial of degree n. If both P_n and P'_n contain a factor $(x - \alpha)^m$, $m < n$, then P_n contains the factor $(x - \alpha)^{m+1}$.*

We also need an expression for $P^{(q)}_{n+1}(k, x)$. By repeatedly differentiating (2) and eliminating $P'_n(k, x)$ we get,

$$P^{(q)}_{n+1}(k, x) = \prod_{i=0}^{q-1} (n + 1 - i)\left[N' + \frac{2k - n + i}{2} D''\right]P_{n-q+1}(k, x), \quad q = 1, 2, \cdots (n + 1). \tag{4}$$

Theorem I_1. If D is a perfect square, D' is not a factor of P_{n+1} (k, x), $n =$ 0, 1, 2, $\cdots$

Proof: Assume D' to be a factor of P_{n+1}. From (1), D' is either a factor of P_n or of $N + (k - n) D'$. But D' is not a factor of $N + (k - n) D'$ as this implies that D' is a factor of N contrary to hypothesis on (R) that D and N have no factor in common. Thus, D' is a factor of P_n, and by a repetition of the reasoning a factor finally of P_1, which as it was just pointed out, is impossible.

Theorem I_2. Set $D = (\alpha_1 x + \beta_1)(\alpha_2 x + \beta_2)$, D not a perfect square. If $\alpha_i x + \beta_i$, $i = 1$ or 2, is a factor of P_n, then $(\alpha_i x + \beta_i)^q$ is a factor of P_{n+q-1}, $q = 1, 2, 3, \cdots$

Proof: From (1), $\alpha_i x + \beta_i$ being a factor of P_n and D, is also a factor of P_{n+1}. From (2), $\alpha_i x + \beta_i$ is a factor of P'_{n+1}. From Lemma (1) it follows that $(\alpha_i x + \beta_i)^2$ is a factor of P_{n+1}. Continued repetition of the reasoning establishes the theorem.

Corollary. If both $\alpha_1 x + \beta_1$ and $\alpha_2 x + \beta_2$ are factors of P_n, then D^q is a factor of P_{n+q-1}.

Theorem I_3. Assume D of the same form as in Theorem I_2. If $\alpha_i x + \beta_i$, $i = 1$ or 2, is a factor of P_{n+1} and no higher power of $\alpha_i x + \beta_i$ is such a factor then $\alpha_i x + \beta_i$ is a factor of $N + (k - n)D'$.

Proof: From (1), $\alpha_i x + \beta_i$ being a factor of P_{n+1} and of D is also a factor of either $N + (k - n)D'$ or of P_n. But $\alpha_i x + \beta_i$ a factor of P_n requires, from I_2, that $(\alpha_i x + \beta_i)^2$ be a factor of P_{n+1} contrary to hypothesis. Thus, $\alpha_i x + \beta_i$ is a factor of $N + (k - n)D'$.

Corollary. If $(\alpha_1 x + \beta_1)(\alpha_2 x + \beta_2)$, $(\alpha_1, \alpha_2 \neq 0)$, *is a factor of* P_{n+1} *and no higher power of either* $\alpha_1 x + \beta_1$ *or* $\alpha_2 x + \beta_2$ *is contained in* P_{n+1} *then* $N + (k - n) D' \equiv 0$. For from I_3, $N + (k - n)D'$ contains $(\alpha_1 x + \beta_1)(\alpha_2 x + \beta_2)$ as a factor which implies $N + (k - n)D'$, being linear, vanishes identically.

Theorem I_4. *If* $(\alpha_i x + \beta_i)^q$ *and no higher power of* $\alpha_i x + \beta_i$ *is a factor of* P_{n+q-1} *then* $\alpha_i x + \beta_i$ *and no higher power of* $\alpha_i x + \beta_i$ *is a factor of* P_n.

Proof: Let us write,

(A) $P_{n+q-1} = (\alpha_i x + \beta_i)^q \phi_{n-1}$, $\phi_{n-1} \equiv$ a polynomial of degree $\leq n - 1$ which does not contain the factor $\alpha_i x + \beta_i$. Taking the $(q - 1)^{\text{st.}}$ derivative of (A) by Leibnitz Theorem, we get,

$$\text{(B)} \qquad P_{n+q-1}^{(q-1)} = \sum_{i=0}^{q-1} \binom{q-1}{i} \frac{d^i}{dx^i} (\alpha_i x + \beta_i)^q \frac{d^{q-1-i}}{dx^{q-1-i}} \phi_{n-1}.$$

On setting $q = q - 1$ in (4) there results,

$$\text{(C)} \quad P_{n+q-1}^{(q-1)} = \prod_{i=0}^{q-2} (n + q - 1 - i) \left[N' + \frac{2k - n - q + i + 2}{2} D'' \right] P_n.$$

From (B) we see that $\alpha_i x + \beta_i$ is a factor of $P_{n+q-1}^{(q-1)}$. No higher power of $\alpha_i x + \beta_i$ is such a factor. From (C) our theorem now follows.

Corollary (1). *Under the hypotheses of Theorem* I_4, $\alpha_i x + \beta_i$ *is a factor of* $N + (k - n + 1)D'$. This follows at once from I_4 and I_3.

Corollary (2). *If* $D^q = (\alpha_1 x + \beta_1)^q (\alpha_2 x + \beta_2)^q$, $(\alpha_1, \alpha_2 \neq 0)$, *is a factor of* P_{n+q-1} *and no higher powers of either* $\alpha_1 x + \beta_1$ *or* $\alpha_2 x + \beta_2$ *are factors, then* $N + (k - n + 1)D' \equiv 0$. For the linear expression $N + (k - n + 1)D'$ contains, from Corollary (1), the quadratic factor $(\alpha_1 x + \beta_1)(\alpha_2 x + \beta_2)$.

The following lemma can be easily established and is given without proof.

Lemma (2). *Assume D of the same form as in Theorem* I_2. *Then there is only one value of s for which* $N + sD'$ *contains* $\alpha_i x + \beta_i$ *as a factor.*

Theorem I_5. *Assume D of the same form as in Theorem* I_2. *If* $N + (k - n)D'$ *contains* $\alpha_i x + \beta_i$, $i = 1$ *or* 2, *as a factor, then* P_{n+1} *contains* $\alpha_i x + \beta_i$ *and no higher power of* $\alpha_i x + \beta_i$ *as a factor.*

Proof: From (1) we see that P_{n+1} contains $\alpha_i x + \beta_i$ at least to the first power as a factor. Again from (1), if P_{n+1} contains a higher power of $\alpha_i x + \beta_i$ as a factor, this means that both P_n and P'_n contain $\alpha_i x + \beta_i$ at least to the first power as a factor and from Lemma (1) it follows that P_n contains $\alpha_i x + \beta_i$ at least to the second power as a factor. By corollary (1) from Theorem I_4 it follows that $\alpha_i x + \beta_i$ is a factor of $N + (k - n_1)D'$ for $n_1 < n$, contrary to Lemma (2).

Theorem I_6. *If* $\alpha_1 x + \beta_1$ *and* $\alpha_2 x + \beta_2$ *are factors of* $N + (k - n_1)D'$ *and* $N + (k - n_2)D'$ *respectively,* $(\alpha_1, \alpha_2 \neq 0)$, *then* $P_\mu \equiv 0$, $\mu > n_1 + n_2$.

Proof: From Theorems I_5 and I_2 we see that $(\alpha_1 x + \beta_1)^{n_2} (\alpha_2 x + \beta_2)^{n_1}$, of degree $n_1 + n_2$, is a factor of $P_{n_2+n_1}$, of degree $n_2 + n_1$ at most. Similarly,

$(\alpha_1 x + \beta_1)^{n_2+1} (\alpha_2 x + \beta_2)^{n_1+1}$, of degree $n_2 + n_1 + 2$, is a factor of $P_{n_2+n_1+1}$, of degree $n_2 + n_1 + 1$ at most. This implies $P_{n_2+n_1+1} \equiv 0$. Hence, $P_\mu \equiv 0$, $\mu > n_1 + n_2$. In fact, (1) shows that $P_\mu \equiv 0$ implies $P_\nu \equiv 0, \nu > \mu$.

Theorem I_7. Assume D of the same form as in Theorem I_2. Then $P_{n+1} \equiv 0$, $P_n \not\equiv 0$, implies either $N + (k - m)D' \equiv 0$, $m \leq n$, or there exist two values of m, (m_1, m_2), such that $N + (k - m_1)D'$, $N + (k - m_2)D'$ contain as factors $\alpha_1 x + \beta_1$ and $\alpha_2 x + \beta_2$ respectively, $(m_1, m_2 \leq n)$.

Proof: Setting $P_{n+1} \equiv 0$ in (1) gives,

$$(1^0)\ [N + (k - n)D']\, P_n + DP_n' \equiv 0.$$

If $P_n \equiv$ const., 1^0 shows that $N + (k - n)D' \equiv 0$ and our theorem is verified. Suppose $P_n \not\equiv$ const. We get from (1^0),

$$P_n' = -\frac{[N + (k - n)D']P_n}{D}.$$

Thus, D is a factor of the numerator, and our theorem now follows from Corollaries (1) and (2) of Theorem I_4.

Theorem I_8. If $N + (k - m)D' \not\equiv 0$, $m = 1, 2, \cdots n$, and if $N + (k - m)D'$ contains neither $\alpha_1 x + \beta_1$, nor $\alpha_2 x + \beta_2$ as factors, then P_{n+1} and D have no factors in common. This follows at once from Theorems I_2 and I_4 which constitute a necessary and sufficient condition that P_n and D have factors in common.

Theorem I_9. If $N \equiv$ const. and if D is linear, all P_n are constants, $n = 1, 2, 3, \cdots$. This follows directly from (2).

Theorem I_{10}. If $N' + \frac{2k - m}{2} D'' \not\equiv 0$, $m = 1, 2, \cdots (n - 1)$, all zeros of P_n which are not zeros of D are simple.

Proof: Suppose P_n has a multiple zero $x = \alpha$ which is not a zero of D. Then (1) shows that α is a zero of P_{n+1}. From (2), α is a zero of P_{n+1}'. From Lemma (1), α is at least a double zero of P_{n+1}. Furthermore, (3) shows that α being a double zero of P_n and of P_{n+1} is also a double zero of P_{n-1}. By a continued application of (3), it follows that α is a double zero of P_1 which is impossible since P_1 is of degree ≤ 1.

II. Concerning the Zeros of $P_n(k, x)$

The polynomials $P_n(k, x)$ are defined by Hildebrandt[3] as follows: $P_n(k, x) = \frac{1}{y} D^{n-k} \frac{d^n}{dx^n} D^k y$ where y is a non-identically vanishing solution of the differential equation

$$\frac{1}{y}\frac{dy}{dx} = \frac{a_0 + a_1 x}{b_0 + b_1 x + b_2 x^2} \equiv \frac{N}{D}.$$

[3] L.c. pp. 400–401.

The Jacobi Polynomials are defined as follows:

$$J_n(x, \alpha, \beta) = x^{1-\alpha}(1-x)^{1-\beta}\frac{d^n}{dx^n}[x^{n+\alpha-1}(1-x)^{n+\beta-1}], \ \alpha, \beta$$

real. It follows that $J_n(x, \alpha, \beta)$ is a special type of $P_n(k, x)$ with $N \equiv (-\beta-\alpha)x + \alpha$, $D \equiv x(1-x)$, $n = k+1$, whence,

$$N' = -\beta-\alpha, \qquad D' = 1-2x, \qquad D'' = -2; \qquad D(0) = D(1) = 0,$$

$$P_1(k, x) \equiv N + kD' = 0 \text{ for}$$

$$x = \frac{\alpha+k}{\alpha+\beta+2k}, \qquad P_1'(k, x) = -\beta-\alpha-2k.$$

In determining the number and location of the real zeros of the Jacobi Polynomials we employ the following notations:

$P_i(k, x) = 0$ for $x = \alpha_{i,k,j}$, $\quad i = 1, 2, \cdots k+1; k = 0, 1, 2, \cdots; j = 1, 2, \cdots i.$

$\alpha_{i,k,j} \leq \alpha_{i+1,k,j}$

$$\theta = N' + \frac{2k-n}{2}D'' = -\beta-\alpha-2k+n, \ \ n = 1, 2, \cdots k,$$

$$\mu = [N + (k-n)D']_{x=0} = \alpha + (k-n),$$

$$\nu = [N + (k-n)D']_{x=1} = -\beta - (k-n).$$

We proceed to determine the number of real zeros of the Jacobi Polynomials on the intervals $(-\infty, 0)$, $(0, 1)$, $(1, \infty)$ into which the zeros of D divide the x axis.[4] The proofs proceed by mathematical induction. We first determine the location of the real zeros of $P_n(k, x)$, $n = 1, 2, \cdots k+1$, by successive applications of (1) and (2). We then use the relation $P_{k+1}(k, x) \equiv J_{k+1}(x, \alpha, \beta)$.

Several cases concerning possible values of α and β should be considered. In order to bring out the method of procedure only two such cases will be fully discussed here. The results for other possible cases will be merely listed.

A_1 : $\alpha < 0, \beta < 0, |\alpha| < |\beta|, \alpha, \beta, \alpha+\beta$ *not integers.*

Let k_1 be the greatest integer contained in α,
" k_2 " " " " " " β,
" k_3 be the greatest integral value of k for which $\alpha+\beta+2k < 0$. Then

$$0 \leq k_1 \leq k_3 \leq k_2.$$

[4] In the case $\alpha, \beta > 0$ these zeros all lie, as is known, on $(0, 1)$.

$A_{11}: 0 \leq k \leq k_1$. *We then have* $\theta > 0$, $\mu < 0$, $\nu > 0$, $0 < \alpha_{1k1} < 1$, $P_1' > 0$. *Then* $J_{k+1}(x, \alpha, \beta)$ *has* $\dfrac{(1)^k + (-1)^k}{2}$ *zeros in* 0, 1. *These are the only real zeros.*

Proof: Consider first $P_1(k, x)$. Its only zero is at $\alpha_{1,k,1}$, where $0 < \alpha_{1,k,1} < 1$. Furthermore, $P_1' > 0$. Also $P_1 > 0$ for $x > \alpha_{1,k,1}$ and < 0 for $x < \alpha_{1,k,1}$. From (1) we see that $P_2(k, \alpha_{1,k,1}) > 0$, (since $P_1(k, \alpha_{1,k,1}) = 0$, $D(\alpha_{1,k,1}) > 0$ and $P_1' > 0$). From (2) it follows that $P_2'(k, x) < 0$ for $x < \alpha_{1,k,1}$, $P_2'(k, \alpha_{1,k,1}) = 0$, $P_2'(k, x) > 0$ for $x > \alpha_{1,k,1}$. These conclusions follow from remarks concerning the sign of θ, the fact that $P_1(k, \alpha_{1,k,1}) = 0$, and from remarks concerning the sign of P_1 to the left and to the right of $x = \alpha_{1,k,1}$. Thus, $P_2(k, x) > 0$ for all real x and hence has no real zeros. By employing (2), it is now evident that $P_3'(k, x) > 0$. From (1) and remarks concerning μ and ν we see that $P_3(k, 0) < 0$ and $P_3(k, 1) > 0$. Thus $P_3(k, x)$ has a single real zero $\alpha_{3,k,1}$, $0 < \alpha_{3,k,1} < 1$. The reasoning from P_3 to P_4 is analogous to that from P_1 to P_2. By continuing this procedure we finally conclude that $P_{k+1}(k, x)$, $(\equiv J_{k+1}(x, \alpha, \beta)$, has but one real zero, (in 0, 1), if k is even and no real zeros if k is odd.

A_{12}: $k_1 < k \leq k_3$. *Set* $k = k_1 + q$, $q = 1, 2, \cdots, k_3 - k_1$. *Here* $\theta > 0$, $\mu > 0, n = 1, 2, \cdots q - 1, \mu < 0, n = q, q + 1, \cdots, q + k_1$. $\nu > 0$, $\alpha_{1,k,1} < 0$, $P_1'(k, x) > 0$. $J_{k_1} + q + 1$ (x, α, β) *has* q *distinct zeros in* $(-\infty, 0)$ *and* $\dfrac{(1)^{k_1} + (-1)^{k_1}}{2}$ *zeros in* 0, 1. *These are the only real zeros.*

Proof: First consider the sequence $P_n(k, x)$ $n = 1, 2, \cdots q$, since the conditions on θ, μ, and ν do not change over this range of n. Now $P_1(k, \alpha_{1,k,1}) = 0$, $\alpha_{1,k,1} < 0$. Furthermore since $P_1' > 0$ we have $P_1 > 0$ for $x > \alpha_{1,k,1}$ and < 0 for $x < \alpha_{1,k,1}$. Pass now to $P_2(k, x)$. Since $D(\alpha_{1,k,1}) < 0$ and $P_1'(k, \alpha_{1,k,1}) > 0$, we see from (1) that $P_2(k, \alpha_{1,k,1}) < 0$. Moreover (2) shows $P_2'(k, \alpha_{1,k,1}) = 0$, $P_2'(k, x) < 0$ for $x < \alpha_{1,k,1}$ and > 0 for $x > \alpha_{1,k,1}$. Thus $P_2(k, x) < 0$ and a relative minimum at $x = \alpha_{1,k,1}$. Since $|P_2(k, \pm\infty)| = \infty$, we see that $P_2(k, x)$ has two real zeros of which the left most, $\alpha_{2,k,1}$, is in $(-\infty, 0)$. Again $\mu > 0$ together with (1) assures $P_2(k, 0) > 0$. Thus $\alpha_{2,k,2}$ is in $(\alpha_{1,k,1}, 0)$, hence in $(-\infty, 0)$. By continuing this reasoning on the successive $P_n(k, x)$, $n = 1, 2, \cdots q$, we conclude that $P_q(k, x)$ has q zeros in $-\infty$, 0 and $P_q'(k, \alpha_{q,k,1}) < 0$.

Next, consider the sequence $P_n(k, x)$, $n = q + 1, q + 2, \cdots q + k_1 + 1$. Over this range of n we have $\theta > 0$, $\mu < 0$, $\nu > 0$. From what has just been shown, $P_q(k, \alpha_{q,k,i}) = 0$, $-\infty < \alpha_{q,k,i} < 0$, $i = 1, 2, \cdots q$. Also $P_q'(k, \alpha_{q,k,i})$, $i = 1, 2, \cdots q$, is alternately negative and positive. Suppose q odd, (similar reasoning holds for q even). Thus, we suppose $P_q'(k, \alpha_{q,k,1}) < 0$, $P_q'(k, \alpha_{q,k,q}) < 0$, $P_q(k, x) > 0$ for $x < \alpha_{q,k,1}$ and < 0 for $x > \alpha_{q,k,q}$. (1) shows $P_{q+1}(k, \alpha_{q,k,i})$, $i = 1, 2, \cdots q$, to be alternately positive and negative. Thus, the zeros $\alpha_{q,k,i}$ are separated by $q - 1$ zeros of $P_{q+1}(k, x)$. Since from (1), $P_{q+1}(k, \alpha_{q,k,1}) > 0$ and from (2) $P_{q+1}'(k, x) > 0$ for $x < \alpha_{q,k,1}$, there exists a zero $\alpha_{q+1,k,1}$ in $(-\infty, \alpha_{q,k,1})$. Thus far, we have established the existence of q zeros of $P_{q+1}(k, x)$ in $(-\infty, 0)$. q being odd, we have from (1), $P_{q+1}(k, \alpha_{q,k,q}) > 0$. Also from (2),

$P'_{q+1}(k, x) < 0$ for $x > \alpha_{q,k,q}$. Again from (1) and assumptions regarding μ and ν it follows that $P_{q+1}(k, 0) > 0$, $P_{q+1}(k, 1) < 0$. Thus, $P_{q+1}(k, x)$ has a zero $\alpha_{q+1,k,q+1}$ in (0, 1). There being no extrema for $P_{q+1}(k, x)$ other than the $\alpha_{q,k,i}$, $i = 1, 2, \cdots q$, (as (2) shows), we have thus proved that $P_{q+1}(k, x)$ has q distinct zeros in $(-\infty, 0)$ and a single zero in (0, 1). Reasoning similarly from $P_{q+1}(k, x)$ to $P_{q+2}(k, x)$ we establish the existence of q distinct zeros $\alpha_{q+2,k,i}$, $i = 1, 2, \cdots q$, in $(-\infty, 0)$ with $\alpha_{q+2,k,1}$ in $(-\infty, \alpha_{q+1,k,1})$ and $\alpha_{q+2,k,i}$, $i = 2, 3, \cdots q$, separating $\alpha_{q+1,k,i}$, $i = 1, 2, \cdots q$. From (1) we see that $P_{q+2}(k, \alpha_{q+1,k,q}) < 0$ and $P_{q+2}(k, \alpha_{q+1,k,q+1}) < 0$. The only extrema of $P_{q+2}(k, x)$, (as (2) shows), are located at $\alpha_{q+1,k,i}$, $i = 1, 2, \cdots q + 1$. Again, by (2), $P'_{q+2}(k, x) < 0$ for $x > \alpha_{q+1,k,q+1}$; hence there can be no real zeros of P_{q+2} except the q zeros in $(-\infty, 0)$ already found. The reasoning from P_{q+2} to P_{q+3} is similar to that from P_q to P_{q+1}. Thus, $P_{q+k_1+1} \equiv J_{k_1+q+1}$ has q distinct zeros in $(-\infty, 0)$, together with one zero in (0, 1) for k_1 even. For k_1 odd, there are q distinct zeros in $(-\infty, 0)$ only. The results are the same whether q is odd or even.

The results for the remaining sub-cases under case A_1 are given in the table which follows. For completeness, the results for cases A_{11} and A_{12} are included in the tabulation. A few words of explanation are necessary to clarify the conditions under which the various sub-cases in the table occur. Let $|\alpha| = k_1 + q$, $|\beta| = k_2 + h$, $h, q < 1$. If $q + h < 1$, then $|\alpha + \beta| = k_1 + k_2$ and we have either,

A_{131} : $k_1 + k_2$ *even*, $2k_3 = k_1 + k_2 \equiv k_3 - k_1 = k_2 - k_3$.

A_{132} : $k_1 + k_2$ *odd*, $2k_3 = k_1 + k_2 - 1 \equiv k_3 - k_1 = k_2 - k_3 - 1$.

Again if $1 < q + h < 2$, then $|\alpha + \beta| = k_1 + k_2 + 1$ and we have either,

A_{133} : $k_1 + k_2 + 1$ *even*, $2k_3 = k_1 + k_2 + 1 \equiv k_3 - k_1 = k_2 - k_3 + 1$.

A_{134} : $k_1 + k_2 + 1$ *odd*, $2k_3 = k_1 + k_2 \equiv k_3 - k_1 = k_2 - k_3$

In cases A_{141} and A_{151} we assume $|\alpha + \beta| = k_1 + k_2 + p$, $p < 1$, while in cases A_{142} and A_{152}, $|\alpha + \beta| = k_1 + k_2 + p$, $1 < p < 2$. The complete results for case A_1 follow. (See page 213.)

A_2 : $\alpha < 0, \beta < 0, |\alpha| < |\beta|$, α, β *not integers*, $\alpha + \beta =$ *integer*. Define k_1, k_2, k_3 as in A_1. Then $0 \leq k_1 \leq k_3 \leq k_2$. In Case A_{21}, $\beta + \alpha$ is odd while in Case A_{22}, $\beta + \alpha$ is even. (See page 214.)

A_3 : $\alpha < 0, \beta < 0, \alpha = -k_1$, *integer*, β *not an integer*, $|\alpha| < |\beta|$. Define k_1, k_2, k_3 as in A_1. Then $0 \leq k_1 \leq k_3 \leq k_2$. There are two sub-cases, A_{31} : the greatest integral value of $\alpha + \beta$ is odd, A_{32} : this integral value is even. (See page 215.)

A_4 : $\alpha < 0, \beta < 0$, α *not an integer*, $\beta = -k_1$, *integer*, $|\alpha| < |\beta|$. Define k_1, k_2, k_3 as in A_1. Then $0 \leq k_1 \leq k_3 \leq k_2$. There are two sub-cases, A_{41} : the integral part of $\alpha + \beta$ is odd, A_{42} : this integral value is even. (See page 216).

Cases	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$	Zeros in $(0, 1)$	Zeros in $(1, \infty)$
A_{11}	J_{k+1};	$0 \leq k \leq k_1$;	0;	$\frac{(1)^k + (-1)^k}{2}$;	0
A_{12}	J_{k_1+q+1};	$q = 1, 2, \cdots, k_3 - k_1$;	q;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0
A_{131}, A_{133};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3$;	$k_3 - k_1 - q$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0
A_{132}, A_{134};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3$;	$k_3 - k_1 - q + 1$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	1
A_{141}	J_{k_2+q+1};	$q = 1, 2, \cdots, k_1$;	$\frac{(1)^{q+1} + (-1)^{q+1}}{2}$;	$\frac{(1)^{k_1+q} + (-1)^{k_1+q}}{2}$;	$\frac{(1)^{k_2-k_1+q+1} + (-1)^{k_2-k_1+q+1}}{2}$
A_{142}	J_{k_2+q+1};	$q = 1, 2, \cdots, k_1$;	$\frac{(1)^{q} + (-1)^{q}}{2}$;	$\frac{(1)^{k_1+q} + (-1)^{k_1+q}}{2}$;	$\frac{(1)^{k_2-k_1+q} + (-1)^{k_2-k_1+q}}{2}$
A_{151}	$J_{k_1+k_2+q+1}$;	$q = 1, 2, \cdots$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	$q - 1$;	$\frac{(1)^{k_2} + (-1)^{k_2}}{2}$
A_{152}	$J_{k_1+k_2+2}$;		$\frac{(1)^{k_1+1} + (-1)^{k_1+1}}{2}$;	0;	$\frac{(1)^{k_2+1} + (-1)^{k_2+1}}{2}$
	$J_{k_1+k_2+q+1}$;	$q = 2, 3, \cdots$;	Same zeros as in A_{151} for corresponding values of q.		

Cases	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$	Zeros in $(0, 1)$	Zeros in $(1, \infty)$
A_{211}, A_{221}	J_{k+1};	$0 \leq k \leq k_1$	0;	$\frac{(1)^k + (-1)^k}{2}$;	0
A_{212}, A_{222}	J_{k_1+q+1};	$q = 1, 2, \cdots, k_3 - k_1$;	q;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0
A_{213}	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3$;	$k_3 - k_1 - q$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0
A_{223}	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3$;	$k_3 - k_1 - q + 1$;	$\frac{(1)^{k_1+1} + (-1)^{k_1+1}}{2}$;	0
A_{214}, A_{224}	J_{k_2+q+1};	$q = 1, 2, \cdots k_1$;	0;	$\frac{(1)^{k_1-q} + (-1)^{k_1-q}}{2}$;	0
A_{215}	$J_{k_1+k_2+2}$;	J = const, > 0, k_1 odd. J = const, < 0, k_1 even.			
	$J_{k_1+k_2+q+2}$;	$q = 1, 2, \cdots$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	q;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$
A_{225}	$J_{k_1+k_2+2}$;	J = const, > 0, k_1 odd. = const, < 0, k_1 even.			
	$J_{k_1+k_2+q+2}$;	$q = 1, 2, \cdots$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	$q - 1$;	$\frac{(1)^{k_1+1} + (-1)^{k_1+1}}{2}$

Cases	Polynomial	Range of Sub-Script	Zeros in			
			$(-\infty, 0)$	$x = 0$	$(0, 1)$	$(1, \infty)$
A_{311}, A_{321};	J_{k+1};	$0 \leq k < k_1$;	0;	0;	$\frac{(1)^k + (-1)^k}{2}$;	0
A_{312}, A_{322};	J_{k_1+q+1};	$q = 0, 1, \cdots, k_3 - k_1$;	q;	$k_1 + 1$;	0;	0
A_{313};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3$;	$k_3 - k_1 - q + 1$;	$k_1 + 1$;	0;	1
A_{323};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3$;	$k_2 - k_1 - q$;	$k_1 + 1$;	0;	0
A_{314};	J_{k_2+q+1};	$q = 1, 2, \cdots, k_1$;	0;	$k_1 + 1$;	0;	$\frac{(1)^q + (-1)^q}{2}$
A_{324};	J_{k_2+q+1};	$q = 1, 2, \cdots, k_1$;	0;	$k_1 + 1$;	0;	$\frac{(1)^{q+1} + (-1)^{q+1}}{2}$
A_{315};	$J_{k_1+k_2+q+1}$;	$q = 1, 2, 3, \cdots$;	0;	$k_1 + 1$;	$q - 1$;	$\frac{(1)^{k_1+1} + (-1)^{k_1+1}}{2}$
A_{325};	$J_{k_1+k_2+q+1}$;	$q = 1, 2, 3, \cdots$;	0;	$k_1 + 1$;	$q - 1$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$

Cases	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$	Zeros in $(0, 1)$	Zeros in $x = 1$	Zeros in $(1, \infty)$
A_{411}, A_{421};	J_{k+1};	$0 \leq k \leq k_1$;	0;	$\frac{(1)^k + (-1)^k}{2}$;	0;	0
A_{412}, A_{422};	J_{k_1+q+1};	$q = 1, 2, \cdots, k_3 - k_1$;	q;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0;	0
A_{413};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3 - 1$;	$k_3 - k_1 - q$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0;	1
A_{423};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3 - 1$;	$k_3 - k_1 - q$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	0;	0
A_{414}, A_{424};	J_{k_2+q+1};	$q = 0, 1, 2, \cdots, k_1$;	$\frac{(1)^{q+1} + (-1)^{q+1}}{2}$;	0;	$k_2 + 1$;	0
A_{415}, A_{425};	$J_{k_1+k_2+q+1}$;	$q = 1, 2, 3, \cdots$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$;	$q - 1$;	$k_2 + 1$;	0

A_5 : $\alpha < 0, \beta < 0, |\alpha| < |\beta|, \alpha = -k_1$ *integer*, $\beta = -k_2$ *integer*. Define k_1, k_2, k_3 as in A_1. In cases A_{51} and A_{52}, $\alpha + \beta$ is odd and even respectively.

Cases	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$	Zeros in $x = 0$	Zeros in $(0, 1)$
A_{511}, A_{521};	J_{k+1};	$0 \leq k < k_1$;	0;	0;	$\frac{(1)^k + (-1)^k}{2}$;
A_{512}, A_{522};	J_{k_1+q+1};	$q = 0, 1, 2, \cdots, k_3 - k_1$;	q;	$k_1 + 1$;	0
A_{513};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3 - 1$;	$k_3 - k_1 - q$;	$k_1 + 1$;	0
A_{523};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_2 - k_3 - 1$;	$k_3 - k_1 - q + 1$;	$k_1 + 1$;	0
A_{514}, A_{524};	$J_{k_2+q+1} \equiv 0$;	$q = 0, 1, 2, \cdots k_1$.			
A_{515}, A_{525};	$J_{k_1+k_2+q+1} \equiv 0$;	$q = 1, 2, 3, \cdots$			

If assumptions are identical with those of A_5 except $|\alpha| = |\beta|$, then for $0 \leq k < k_1$, the results agree with A_{511} and $J_{k_1+q+1} \equiv 0, q = 0, 1, 2, \cdots$.

A_6 : $\alpha > 0, \beta < 0, |\alpha| > |\beta|, \beta$ *not an integer*. Let k_1 be the largest integer in β.

Case	Polynomial	Range of Sub-Script	Zeros in $(0, 1)$	Zeros in $(1, \infty)$
A_{61}	J_{k+1}	$0 \leq k < k_1$	0	$\frac{(1)^k + (-1)^k}{2}$
A_{62}	J_{k_1+q+1}	$q = 1, 2, 3, \cdots$	q	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$

A_7 : *Same assumptions as in* A_6 *except* $\beta = -k_1$, *integer*.

Case	Polynomial	Range of Sub-Script	Zeros in $(0, 1)$	Zeros in $x = 1$	Zeros in $(1, \infty)$
A_{71}	J_{k+1}	$0 \leq k \leq k_1 - 1$	0	0	$\frac{(1)^k + (-1)^k}{2}$
A_{72}	J_{k_1+q+1}	$q = 0, 1, 2, \cdots$	q	$k_1 + 1$	0

A_8 : $\alpha > 0, \beta < 0, |\alpha| = |\beta|$. $J_1 = \alpha$ and results for J_n, $n > 1$ are identical with those in A_7 and A_6 respectively according as β is or is not an integer.

A_9 : $\alpha > 0, \beta < 0, |\alpha| < |\beta|$; $\beta, \alpha + \beta$, *not integers*.

Let k_1 be the greatest integer in $\alpha + \beta$.
" k_2 " " " " " β.
" k_3 " " " " for which $\alpha + \beta + 2k < 0$.

Then $0 \leq k_3 \leq k_1 \leq k_2$.

Case	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$	$(0, 1)$	$(1, \infty)$
A_{91};	J_{k+1};	$0 \leq k \leq k_3$;	$k + 1$;	0;	0
A_{921};	J_{k_3+q+1};	$q = 1, 2, \cdots, k_3$; k_1 even;	$k_3 - q + 1$;	0;	0
A_{922};	J_{k_3+q+1};	$q = 1, 2, \cdots, (k_3 + 1)$; k_1 odd;	$k_3 - q + 2$;	0;	1
A_{93};	J_{k_1+q+1};	$q = 1, 2, \cdots, (k_2 - k_1)$;	0;	0;	$\frac{(1)^{k_1+q} + (-1)^{k_1+q}}{2}$
A_{94};	J_{k_2+q+1};	$q = 1, 2, 3, \cdots$;	0;	q;	$\frac{(1)^{k_2} + (-1)^{k_2}}{2}$

A_{10} : *Same assumptions as in* A_9 *but now* $|\alpha| = |\beta|$. Then $k_1 = k_3 = 0$, $J_1 = \alpha$, and results for J_n, $n > 1$ are the same as in A_{93} and A_{94}.

A_{11} : *Same assumptions as in* A_9 *except* $\beta = -k_2$, *integer.*

Case	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$	$(0, 1)$	$x = 1$	$(1, \infty)$
$A_{11,1}$	Same as A_{91}					
$A_{11,2}$	Same as A_{92}					
$A_{11,3}$	Same as A_{93}					
$A_{11,4}$	J_{k_2+q+1};	$q = 1, 2, 3, \cdots$;	0;	q;	$k_2 + 1$;	0

A_{12} : $\alpha > 0$, $\beta < 0$, $|\alpha| < |\beta|$, β *not an integer.* $\alpha + \beta =$ *odd integer.* Define k_1, k_2, k_3 as in A_9.

A_{13} : *Same assumptions as in* A_{12} *except* $\alpha + \beta =$ *even integer.*

Cases	Polynomial	Range of Sub-Script	Zeros in $(-\infty, 0)$
$A_{12,1}$, $A_{13,1}$;	Same as A_{91}		
$A_{12,2}$;	$\begin{cases} J_{k_3+q+1}; \\ J_{2k_3+3} = \text{const.} > 0; \end{cases}$	$q = 1, 2, \cdots, k_3$;	$k_3 - q + 1$
$A_{13,2}$;	$\begin{cases} J_{k_3+q+1}; \\ J_{2k_3+3} = \text{const.} > 0; \end{cases}$	$q = 1, 2, \cdots, k_3 + 1$;	$k_3 - q + 2$
$A_{12,3}$, $A_{13,3}$;	Same as A_{93}		
$A_{12,4}$, $A_{13,4}$;	Same as A_{94}		

A_{14} : *Same assumptions as in* A_{12}, *except* $\beta = -k_2$ *integer.* Cases $A_{14,1}$, $A_{14,2}$ and $A_{14,3}$ have the same results as $A_{12,1}$, $A_{12,2}$, and $A_{12,3}$ respectively. $A_{14,4}$ has the same results as $A_{11,4}$.

A_{15} : *Same assumptions as* A_{13} *except* $\beta = -k_2$, *integer.* Cases $A_{15,1}$, $A_{15,2}$, and $A_{15,3}$ have the same results as $A_{13,1}$, $A_{13,2}$, and $A_{13,3}$ respectively. $A_{15,4}$ has the same results as $A_{11,4}$.

A_{16} : $\alpha = 0, \beta < 0, \beta$ – *not an integer.*

Let k_1 be the largest integer contained in β.

" k_3 be the largest integer for which $\beta + 2k < 0$.

Case	Polynomial	Range of Sub-Script	Zeros in			
			$(-\infty, 0)$	$x = 0$	$(0, 1)$	$(1, \infty)$
$A_{16,1}$;	J_{k+1};	$0 \leq k \leq k_3$;	k;	1;	0;	0
$A_{16,2}$;	J_{k_3+q+1};	$q = 1, 2, \cdots, k_1 - k_3$;	$k_3 - q$;	1;	0;	0; k_1 even
			$k_3 - q + 1$;	1;	0;	1; k_1 odd
$A_{16,3}$;	J_{k_1+q+1};	$q = 1, 2, 3, \cdots$;	0;	1;	$q - 1$;	$\frac{(1)^{k_1} + (-1)^{k_1}}{2}$

A_{17} : $\alpha = 0, \beta = -k_1$ – *odd integer.* Define k_3 as in A_{16}.

A_{18} : $\alpha = 0, \beta = -k_1$ – *even integer.* Define k_3 as in A_{16}.

Cases	Polynomial	Range of Sub-Script	Zeros in			
			$(-\infty, 0)$	$x = 0$	$(0, 1)$	$x = 1$
$A_{17,1}$, $A_{18,1}$;	Same as $A_{16,1}$					
$A_{17,2}$;	J_{k_3+q+1};	$q = 1, 2, \cdots, k_1 - k_3 - 1$;	$k_3 - q$;	1;	0;	0
$A_{18,2}$;	J_{k_3+q+1};	$q = 1, 2, \cdots, k_3 + 1$;	$k_3 - q + 1$;	1;	0;	0
$A_{17,3}$, $A_{18,3}$;	$\begin{cases} J_{k_1+1} \equiv 0 \\ J_{k_1+q+1}; \end{cases}$	$q = 1, 2, 3, \cdots$;	0;	1;	$q - 1$;	$k_1 + 1$

A_{19} : $\alpha = 0, \beta = 0. \quad J_1 \equiv 0.$

J_{k+1} has $k - 1$ zeros in $(0, 1)$, 1 zero at $x = 0$, 1 zero at $x = 1$, $k = 1, 2, 3, \cdots$.

From the definition of $J_n(x, \alpha, \beta)$ it is readily seen that $J_n(x, \alpha, \beta) \equiv (-1)^n J_n(1 - x, \beta, \alpha)$. Thus, a transformation of x to $1 - x$ interchanges α and β. The interval $(-\infty, 0)$ is transformed into $(1, \infty)$ and vice-versa. The points $x = 0$ and $x = 1$ are interchanged. Consequently, in all previous results we may interchange properly α and β.

In the foregoing results, the only real multiple zeros that can occur are at either $x = 0$ or $x = 1$. In the process of determining the degree of multiplicity of such zeros use was made of Theorem I_2.

Points of Inflexion. By taking (4), setting $k = n$, and replacing N' and D''

by their values for Jacobi polynomials, we get: $P''_{n+1}(n, x) = (n + 1)(n)[\beta + \alpha + n][\beta + \alpha + n + 1]P_{n-1}(n, x)$. From definitions of $P_n(k, x)$ and $J_n(x, \alpha, \beta)$ we easily verify that,

$$P_n(n \pm q, x) \equiv J_n(x, \alpha \pm q + 1, \beta \pm q + 1), \text{ whence,}$$

$$J''_n(x, \alpha, \beta) = (n + 1)(n)[\beta + \alpha + n][\beta + \alpha + n + 1]J_{n-1}(x, \alpha + 2, \beta + 2).$$

We conclude that if neither $\alpha + \beta + n$ nor $\alpha + \beta + n + 1$ vanishes, the points of inflexion of $J_{n+1}(x, \alpha, \beta)$ are at the zeros of odd order of $J_{n-1}(x, \alpha + 2, \beta + 2)$.

The Degree of $J_n(x, \alpha, \beta)$. In analyzing the results of cases A_1 to A_{19} inclusive, it is noted that in some cases the number of real zeros of J_n is less than n. The question naturally arises whether the degree of J_n is n or less, for then we can determine the number of its imaginary zeros. The explicit expression of $J_n(x, \alpha, \beta)$ is known from which the degree of J_n can be found for various α and β. However, the degree of J_n can be found from (4).

Since $J_{n+1}(x, \alpha, \beta) \equiv P_{n+1}(n, x)$, let us replace k by n in (4) and at the same time replace N' and D'' by their values for Jacobi Polynomials. Thus, we get:

$$J^{(q)}_{n+1}(x, \alpha, \beta) = \prod_{i=0}^{q-1}(n + 1 - i)[-\beta - \alpha - n - i]P_{n-q+1}(n, x), \tag{5}$$
$$n = 0, 1, 2, \cdots; q = 0, 1, \cdots, (n + 1).$$

We may establish the following results.

C_1) If $\alpha + \beta$ is not an integer, the degree of $J_{n+1}(x, \alpha, \beta)$ is $n + 1$, $n = 0, 1, 2, \cdots$.

In fact, in order for $J^{(q)}_{n+1}$ to vanish, we see from (5) that either some factor $-\beta - \alpha - n - i$ vanishes or $P_{n-q+1}(n, x)$ vanishes identically. We first show that the latter is not possible. Now $P_1(n, x) \equiv N + nD' \equiv (-\beta - \alpha - 2n)x + \alpha + n \not\equiv 0$ since $\beta + \alpha$ is not an integer. Consequently, if $P_\mu(n, x) \equiv 0$, $\mu > 0$, $\mu \leq n + 1$ there will be a first value of μ, $(\mu = \nu)$, for which $P_\nu(n, x) \equiv 0$ but $P_{\nu-1}(n, x) \not\equiv 0$. By virtue of Theorem I_7 this means that either $N + (n - p)D' \equiv [-\beta - \alpha - 2(n - p)]x + \alpha + n - p \equiv 0$, $p \leq \nu$, or else there exist two values of p, (p_1, p_2), such that $[-\beta - \alpha - 2(n - p_1)]x + \alpha + n - p_1$ and $[-\beta - \alpha - 2(n - p_2)]x + \alpha + n - p_2$ are divisible by x and $1 - x$ respectively, $p_1, p_2 \leq \nu - 1$, $p_1 \neq p_2$. Since, however, $\alpha + \beta$ is not an integer we see that, $[-\beta - \alpha - 2(n - p)]x + \alpha + n - p \not\equiv 0$, n and p being integers. This eliminates the first possibility that $P_\mu(n, x) \equiv 0$, $\mu \leq n + 1$. Again, if, $[-\beta - \alpha - 2(n - p_1)]x + \alpha + n - p_1$ is divisible by x, we have $\alpha + n - p_1 = 0$ or α an integer. For $(\alpha + n - p_2) - [\beta + \alpha + 2(n - p_2)]x \equiv (\alpha + n - p_2)\left[1 - \dfrac{(\alpha + n - p_2) + (\beta + n - p_2)}{(\alpha + n - p_2)}x\right]$ to be divisible by $1 - x$ requires $\beta + n - p_2 = 0$ or β, an integer. α and β are therefore both integers contrary to hypothesis. Thus, in (5), no polynomial $P_{n-q+1}(k, x) \equiv 0$ and $J_{n+1}(x, \alpha, \beta) \not\equiv 0$. Replacing q by $n + 1$ in (5) leads to,

$$J_{n+1}^{(n+1)}(x, \alpha, \beta) = \prod_{i=0}^{n} (n+1-i)[-\beta-\alpha-n-i]\, P_0(n, x), \tag{6}$$
$$n = 0, 1, 2, \cdots.$$

Thus $J_{n+1}^{(n+1)} \not\equiv 0$, (since $P_0(n, x) = 1$ and no factor $-\beta - \alpha - n - i$ can vanish) and the degree of J_{n+1} is precisely $n + 1$. From similar reasoning we prove:

C$_2$) If $\alpha + \beta > 0$ the degree of J_{n+1} is $n + 1$, $n = 0, 1, 2, \cdots$.

C$_3$) If $\alpha + \beta = 0$, then (I) $J_1 = \alpha$ and (II) J_{n+1} is of degree $n + 1$, $n = 1, 2, 3, \cdots$.

C$_4$) If $\alpha + \beta = -M$ – integer, $M > 0$, β, α not integers, then,
 (I) For $n < M$, the degree of J_{n+1} is min. $(n + 1, M - n)$.
 (II) $n = M$, $J_{n+1} \equiv$ const.
 (III) $n > M$, the degree of J_{n+1} is $n + 1$.

C$_5$) If $\alpha + \beta = -M$ – integer, $M > 0$, α, β integers, $\alpha > 0$, $\beta < 0$, then,
 (I) For $n < M$, the degree of J_{n+1} is min. $(n + 1, M - n)$.
 (II) $n = M$, $J_{n+1} \equiv$ const.
 (III) $n > M$, the degree of J_{n+1} is $n + 1$.

C$_6$) If $\alpha + \beta = -M$ – integer, $M > 0$, $\alpha = -k_1$-integer, $\beta = -k_2$-integer, $k_1 < k_2$ then,
 (I) For $n < k_2$, J_{n+1} is of degree $n + 1$.
 (II) $n \geq k_2$, $J_{n+1} \equiv 0$.

C$_7$) If $\alpha + \beta = -M$ – integer, $M > 0$, $\alpha = \beta = -k_1$-integer, then,
 (I) For $n < k_1$, J_{n+1} is of degree $n + 1$,
 (II) $n \geq k_1$, $J_{n+1} \equiv 0$.

The Laguerre Polynomials. These are defined as follows:

$$L_n \equiv L_n(x, \alpha) = x^{1-\alpha} e^x \frac{d^n}{dx^n} [e^{-x} x^{n+\alpha-1}], \qquad n = 0, 1, 2, \cdots;$$

α – real. We see that L_n is a special case of $P_n(k, x)$ with $N \equiv -x + \alpha$, $D \equiv x$, $n = k + 1$. It follows that $\theta = -1$, $\mu = \alpha + k - n$, $\alpha_{1k1} = \alpha + k$, and $P_1'(k, x) = 1$. These can be used in determining the location of the real zeros of L_n, as was done for J_n. The discussion here is somewhat simplified since L_n has but one parameter, α, and the x-axis is divided by the zeros of $D(x)$ into two segments only, namely, $(-\infty, 0)$ and $(0, \infty)$.

The following results are easily obtained.

B$_1$: $\alpha > 0$, $L_n(x, \alpha)$ has n distinct zeros in $(0, \infty)$, $n = 1, 2, 3, \cdots$. This result is well known.

B$_2$: $\alpha = 0$. $L_{n+1}(x, \alpha)$ has n distinct zeros in $(0, \infty)$ and a simple zero at $x = 0$, $n = 0, 1, 2, \cdots$.

B$_3$: $\alpha < 0$, α, *not an integer*. Let k_1 be the largest integer contained in α.

(I) $L_{k+1}(x, \alpha)$ has $\dfrac{(1)^k + (-1)^k}{2}$ zeros in $(-\infty, 0)$, $0 \leq k \leq k_1$,

(II) $L_{k_1+q+1}(x, \alpha)$ has q distinct zeros in $(0, \infty)$ and $\frac{(1)^{k_1} + (-1)^{k_1}}{2}$ zeros in $(-\infty, 0), q = 0, 1, 2, \cdots$.

B_4 : $\alpha < 0$, $\alpha = -k_1$ — *integer.*

(I) $L_{k+1}(x, \alpha)$ has $\frac{(1)^k + (-1)^k}{2}$ zeros in $(-\infty, 0), 0 \leq k \leq k_1$.

(II) $L_{k_1+q+1}(x, \alpha)$ has q distinct zeros in $(0, \infty)$ and a zero of order $k_1 + 1$ at $x = 0, q = 0, 1, 2, \cdots$.

The Degree of $L_n(x, \alpha)$. We show first that here $P\mu(n, x) \not\equiv 0$, $\mu = 1, 2, \cdots n + 1$. By definition, $P_1(n, x) \equiv N + nD' \equiv -x + \alpha + n \not\equiv 0$. Let us rewrite (2) for our present situation thus:

(2°) $P'_\mu(n, x) = -\mu P_{\mu-1}(n, x)$. If, now, $P_\mu(n, x) \equiv 0$, then from (2°) it follows that $P_{\mu-1}(n, x) \equiv 0$. Continuing this reasoning, we finally arrive at a contradiction, namely, $P_1(n, x) \equiv 0$. If in (4) we set $q = n + 1$ and replace N' and D'' by their values we get:

$$L_{n+1}^{(n+1)}(x, \alpha) = (-1)^{n+1}(n + 1)! \qquad P_0(n, x) = (-1)^{n+1}(n + 1)!$$

Hence, L_{n+1} is of degree $n + 1$. Note that this holds regardless of the value of α contrary to what was found for Jacobi Polynomials.

Points of Inflexion. By a procedure analogous to that used for Jacobi Polynomials we can show that the points of inflexion of $L_{n+1}(x, \alpha)$ are located at the zeros of odd order of $L_{n-1}(x, \alpha + 2)$.

The Polynomials $P_n(0, x)$. If we set $k = 0$ in (1), (2), and (3) we obtain the following relationships for $P_n(0, x)$[5] $\equiv P_n(x) \equiv P_n$.

(7) $P_{n+1}(x) = [N - nD'] P_n(x) + DP'_n(x)$.

(8) $P'_{n+1}(x) = (n + 1)\left[N' - \frac{n}{2} D''\right] P_n(x)$.

(9) $P_{n+1}(x) = [N - nD'] P_n(x) + n\left(N' - \frac{n-1}{2} D''\right) DP_{n-1}(x)$.

Theorems I_1 to I_{10} inclusive, with $k = 0$, hold for $P_n(x)$. In addition, the following theorems hold for P_n.

Theorem H_1. *Suppose N linear and* $D(x) > 0$ *for all x. Furthermore, let* $N' - \frac{m}{2} D'' < 0$, $m = 1, 2, 3, \cdots$. *Then* P_n *has n real, distinct zeros which separate the zeros of* P_{n+1}.

Proof: Denote the zeros of P_n by $\alpha_{n,i}$, $i = 1, 2, \cdots n$, $\alpha_{n,i} < \alpha_{n,i+1}$. Suppose $N' > 0$. N being linear has a single zero α_{11}. Furthermore, since $P_1 \equiv N_1$, then $P_1 < 0$ for $x < \alpha_{11}$ and > 0 for $x > \alpha_{11}$. We pass now to P_2. From (7), we see that $P_2(\alpha_{11}) > 0$, (since $D > 0$ and $P'_1 > 0$). Also (8) shows $P'_2(x) > 0$

[5] E. H. Hildebrandt, loc. cit. pp. 399.

for $x < \alpha_{11}$ and < 0 for $x > \alpha_{11}$. This follows from what was noted concerning the sign of P_1 for $x > \alpha_{11}$ and $x < \alpha_{11}$, together with the hypothesis that $N' - \frac{m}{2} D'' < 0$. Thus, there exists a zero of P_2 in $(-\infty, \alpha_{11})$ and a zero in (α_{11}, ∞) and our theorem holds for $n = 1$. Assume that the theorem is true for $n = h$. The sequence $P_h'(\alpha_{h,i})$, $i = 1, 2, \cdots h$, is alternately positive and negative. Since, from (8), the only extrema of P_{h+1} are at $\alpha_{h,i}$, $i = 1, 2, \cdots h$, we conclude that there are $h - 1$ zeros of P_{h+1} separating the $\alpha_{h,i}$, $i = 1, 2, \cdots h$. Since $P_h'(\alpha_{h,1}) > 0$ we conclude that $P_h < 0$ for $x < \alpha_{h,1}$. This fact, combined with (8), shows $P_{h+1}'(x) > 0$ for $x < \alpha_{h,1}$. $P_{h+1}(\alpha_{h,1})$ being positive, it follows that there exists a zero of P_{h+1} in $(-\infty, \alpha_{h,1})$. Similar reasoning establishes the existence of a zero of P_{h+1} in $(\alpha_{h,h}, \infty)$. Our theorem is thus established for $N' > 0$. The case $N' < 0$ can be similarly treated.

Theorem H_2: If $D(x) > 0$ for all x, $D'' < 0$, $N' - \frac{m}{2} D'' < 0$, $N' = 0$, $N \not\equiv 0$, then P_n, $n = 2, 3, \cdots$, has $n - 1$ real, distinct zeros which are separated by the zeros of P_{n-1}.

Proof: Since $P_1 \equiv N = \text{const.}$, we see from (7) that P_2 is linear. The reasoning of Theorem H_1 applies where we now start with P_2.

Theorem H_3: If $D(x) > 0$ for all x, except $x = \beta$, where D has a double zero and if $N' \not\equiv 0$, $N' - \frac{n}{2} D'' < 0$, $n = 1, 2, 3, \cdots$, then P_n has n real, distinct zeros which separate those of P_{n+1}.

Proof: Theorem I_1 with $k = 0$ assures us that P_n and D have no zeros in common. The proof now follows the line of reasoning of Theorem H_1.

Theorem H_4: If $D(x) > 0$ for all x except $x = \beta$ where D has a double zero and if $N' = 0$, $N \not\equiv 0$, $N' - \frac{m}{2} D'' < 0$, $m = 1, 2, 3, \cdots$, then P_n has $n - 1$ real, distinct zeros which separate those of P_{n+1}, $n = 1, 2, 3, \cdots$. This theorem follows from H_3 as did H_2 from H_1.

Points of Inflexion. Setting $k = 0$ in (4) leads to,

$$P_{n+1}'' = (n+1)(n)\left[N' - \frac{n}{2} D''\right]\left[N' - \frac{n-1}{2} D''\right] P_{n-1}.$$

This shows, under the assumptions of Theorems H_1 to H_4 inclusive, that the points of inflexion of P_{n+1} are at the zeros of P_{n-1}.

Hermite Polynomials. Theorem H_1 and statement immediately above concerning points of inflexion apply directly to Hermite Polynomials where $N \equiv -x$ and $D \equiv \sigma^2$.

LEHIGH UNIVERSITY.

THE SIMULTANEOUS COMPUTATION OF GROUPS OF REGRESSION EQUATIONS AND ASSOCIATED MULTIPLE CORRELATION COEFFICIENTS

By Paul S. Dwyer

1. **Introduction.** The need sometimes arises for the prediction of a number of different variables from a given group of so-called fundamental variables. In the work of college prediction, for example, one might desire regression equations predicting certain measures of college achievement (e.g., first semester average, first semester English grade, first semester mathematics grade, number of hours of A received during first semester, etc.) on the basis of a number of other factors (e.g., high school record, score on American Council on Education Psychological Examination, score on some standard English achievement test, score on some standard mathematics achievement test, etc.). It is the purpose of this paper to show how the regression coefficients and the associated multiple correlation coefficients can be obtained simultaneously. The essence of the method is a simple device by which one solution of general normal equations may be made to serve for all cases.

2. **The normal equations.** Let $x_1, x_2, x_3, \cdots x_n$, be the so-called fundamental variables and let x_k be the predicted variable. The normal equations are computed by standard methods which result in one of the three types.

Type I. Normal equations for determining $b_0, b_1, b_2, b_3, \cdots, b_n$.

$$\begin{array}{l}
b_0 n + b_1\Sigma x_1 + b_2\Sigma x_2 + b_3\Sigma x_3 + \cdots\cdots + b_n\Sigma x_n - \Sigma x_k = 0 \\
b_0\Sigma x_1 + b_1\Sigma x_1^2 + b_2\Sigma x_1x_2 + b_3\Sigma x_1x_3 + \cdots\cdots + b_n\Sigma x_1x_n - \Sigma x_1x_k = 0 \\
b_0\Sigma x_2 + b_1\Sigma x_1x_2 + b_2\Sigma x_2^2 + b_3\Sigma x_2x_3 + \cdots\cdots + b_n\Sigma x_2x_n - \Sigma x_2x_k = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
b_0\Sigma x_n + b_1\Sigma x_nx_1 + b_2\Sigma x_nx_2 + b_3\Sigma x_nx_3 + \cdots\cdots + b_n\Sigma x_n^2 - \Sigma x_nx_k = 0
\end{array}$$

Type II. Normal equations for determining $b_1, b_2, b_3, \cdots, b_n$.

$\bar{x}_i = x_i - M_{x_i}$

$$\begin{array}{l}
b_1\Sigma \bar{x}_1^2 + b_2\Sigma \bar{x}_1\bar{x}_2 + b_3\Sigma \bar{x}_1\bar{x}_3 + \cdots\cdots + b_n\Sigma \bar{x}_1\bar{x}_n - \Sigma \bar{x}_1\bar{x}_k = 0 \\
b_1\Sigma \bar{x}_2\bar{x}_1 + b_2\Sigma \bar{x}_2^2 + b_3\Sigma \bar{x}_2\bar{x}_3 + \cdots\cdots + b_n\Sigma \bar{x}_2\bar{x}_n - \Sigma \bar{x}_2\bar{x}_k = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
b_1\Sigma \bar{x}_n\bar{x}_1 + b_2\Sigma \bar{x}_n\bar{x}_2 + b_3\Sigma \bar{x}_n\bar{x}_3 + \cdots\cdots + b_n\Sigma \bar{x}_n^2 - \Sigma \bar{x}_n\bar{x}_k = 0
\end{array}$$

Type III. Normal equations for determining $\beta_1, \beta_2, \beta_3, \cdots, \beta_n$.

$$
\begin{array}{l}
\beta_1 + r_{12}\beta_2 + r_{13}\beta_3 + \cdots\cdots + r_{1n}\beta_n - r_{1k} = 0 \\
r_{21}\beta_1 + \beta_2 + r_{23}\beta_3 + \cdots\cdots + r_{2n}\beta_n - r_{2k} = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
r_{n1}\beta_1 + r_{n2}\beta_2 + r_{n3}\beta_3 + \cdots\cdots + r_{nn}\beta_n - r_{nk} = 0
\end{array}
$$

The three types are special cases of the general

$$
\begin{array}{l}
d_{11}y_1 + d_{12}y_2 + d_{13}y_3 + \cdots + d_{1j}y_j + \cdots + d_{1n}y_n - d_{1k} = 0 \\
d_{21}y_1 + d_{22}y_2 + d_{23}y_3 + \cdots + d_{2j}y_j + \cdots + d_{2n}y_n - d_{2k} = 0 \\
d_{31}y_1 + d_{32}y_2 + d_{33}y_3 + \cdots + d_{3j}y_j + \cdots + d_{3n}y_n - d_{3k} = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
d_{i1}y_1 + d_{i2}y_2 + d_{i3}y_3 + \cdots + d_{ij}y_j + \cdots + d_{in}y_n - d_{ik} = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
d_{n1}y_1 + d_{n2}y_2 + d_{n3}y_3 + \cdots + d_{nj}y_j + \cdots + d_{nn}y_n - d_{nk} = 0
\end{array}
$$

where y_j are the regression coefficients and $d_{ij} = d_{ji}$.

The methods described in this paper are applicable to the general case and hence to each of the three particular types.

In examining the normal equations, it is noticed that the first n terms of each equation are completely determined by the n fundamental variables. The equations, aside from the last terms, are identical no matter what variable is predicted. It is only necessary to devise a technique for separating the contributions of the d_{ik} terms.

3. Solution by determinants. One method utilizes determinants. The value y_j is expressed in terms of a determinant involving a column with entries $d_{1k}, d_{2k}, d_{3k}, \cdots, d_{nk}$. The determinant is expanded in terms of this column.

Specifically, let D be the determinant of the coefficients of the y_j and let D_{ij} be the cofactor of any element d_{ij} of D. Then

$$D = \sum_{i=1}^{n} D_{ij}\, d_{ij}$$

and

$$y_1 = \frac{1}{D}(D_{11}\, d_{1k} + D_{21}\, d_{2k} + D_{31}\, d_{3k} + \cdots + D_{j1}\, d_{jk} + \cdots + D_{n1}\, d_{nk}.)$$

$$y_2 = \frac{1}{D}(D_{12}\, d_{1k} + D_{22}\, d_{2k} + D_{32}\, d_{3k} + \cdots + D_{j2}\, d_{jk} + \cdots + D_{n2}\, d_{nk}.)$$

$$y_i = \frac{1}{D}(D_{1i}\, d_{1k} + D_{2i}\, d_{2k} + D_{3i}\, d_{3k} + \cdots + D_{ji}\, d_{jk} + \cdots + D_{ni}\, d_{nk}.)$$

$$y_n = \frac{1}{D}(D_{1n}\, d_{1k} + D_{2n}\, d_{2k} + D_{3n}\, d_{3k} + \cdots + D_{jn}\, d_{jk} + \cdots + D_{nn}\, d_{nk}.)$$

It is only necessary to compute $\frac{D_{ji}}{D}$ to find the coefficient of d_{jk} in the expansion of y_i.

An illustration is given. The normal equations are

$$\beta_1 + .3300\,\beta_2 + .2100\,\beta_3 - r_{1k} = 0$$

$$.3300\,\beta_1 + \beta_2 - .4800\,\beta_3 - r_{2k} = 0$$

$$.2100\,\beta_1 - .4800\,\beta_2 + \beta_3 - r_{3k} = 0$$

from which at once

$$\beta_1 = \frac{1}{D}(.7696\,r_{1k} - .4308\,r_{2k} - .3684\,r_{3k})$$

$$\beta_2 = \frac{1}{D}(-.4308\,r_{1k} + .9559\,r_{2k} + .5493\,r_{3k})$$

$$\beta_3 = \frac{1}{D}(-.3684\,r_{1k} + .5493\,r_{2k} + .8911\,r_{3k})$$

and also

$$\begin{aligned} D = .550072 &= (1.00)(.7696) + (.33)(-.4308) + (.21)(-.3684) \\ &= (.33)(-.4308) + (\ 1.00)(.9559) + (-.48)(.5493) \\ &= (.21)(-.3684) + (-.48)(.5493) + (\ 1.00)(.8911) \end{aligned}$$

so that

$$\beta_1 = \ \ 1.3991\,r_{1k} - \ .7832\,r_{2k} - \ .6697\,r_{3k}\,.$$

$$\beta_2 = -.7832\,r_{1k} + 1.7378\,r_{2k} + \ .9986\,r_{3k}\,.$$

$$\beta_3 = -.6697\,r_{1k} + \ .9986\,r_{2k} + 1.6200\,r_{3k}\,.$$

It is only necessary to insert any given values r_{1k}, r_{2k}, r_{3k}, to obtain the coefficients of any specific regression equation.

4. **Solutions without determinants.** Theoretically the solution by determinants is excellent but as the number of variables increases the work of computing the n^2 cofactors $\left[\text{or the } \frac{n(n+1)}{2} \textit{ different} \text{ cofactors}\right]$ becomes enormous. We desire a technique for separating the contributions of the last terms when determinants are not used. This can be accomplished by using a separate column for each d_{ik}. Before algebraic manipulation, the value d_{ik} is factored from the column and, after manipulative solution is complete, the multiplication by d_{ik} is carried out.

As an example consider the normal equations

$$\beta_1 + r_{12}\beta_2 - r_{1k} = 0$$

$$r_1\beta_{21} + \beta_2 - r_{2k} = 0$$

where $r_{12} = r_{21} = .3300$. Then the normal equations may be represented by rows (1) and (2) of Table I.

TABLE I

Row	Operation	β_1	β_2	r_{1k}	r_{2k}
(1)		1.0000	.3300	−1.0000	
(2)		.3300	1.0000		−1.0000
(3)	− .3300 times (2)	− .1089	− .3300		.3300
(4)	(1) + (3)	.8911		−1.0000	.3300
(5)	− (4) divided by .8911	−1.0000		1.1222	− .3703
(6)	− .3300 times (5)	.3300		− .3703	.1222
(7)	− (2) + (6)		−1.0000	− .3703	1.1222

The four decimal place solution, whose steps are indicated by (3) (4) (5) (6)(7), is from (5) and (7)

$$\beta_1 = 1.1222\, r_{1k} - .3703\, r_{2k}$$

$$\beta_2 = -.3703\, r_{1k} + 1.1222\, r_{2k}$$

This device may be combined with most of the standard methods of solving normal equations.

5. Combination with Doolittle method. Especially to be recommended is a combination of this device with the Doolittle method which is recognized as a most efficient method of solving normal equations in from five to ten variables [1] [2]. One of the advantages of the Doolittle method is that related multiple regression coefficients may be obtained from the same forward solution, though additional back solutions are necessary [3].

The problem which led to the development of this technique was the simultaneous prediction of scores on various occupations covered by the Strong Vocational Interest Blank from the scores on a few fundamental occupations. A multiple factor analysis revealed that five basic factors account for most of the scores. Five occupational scores, serving as approximations to the five basic factors, were used as the fundamental variables and the other scores were predicted from them.

As an illustration of this prediction technique combined with the Doolittle method, I have selected three test scores as fundamental since the solution based on them shows all the steps of the Doolittle method and is shorter than the five

variable problem. Actually, solution by determinants (section 3) is advised for problems involving three variables. The steps of the Doolittle solution are presented in Table II. The results should be compared with those of the determinant solution of section 3.

The first column indicates the row and the second the description of the algebraic operation. The next three columns are the standard columns of a Doolittle presentation with the conventional elimination of the lower left entries. The next three columns carry through the Doolittle method with the values r_{1k}, r_{2k}, r_{3k} kept in separate columns. The last column is an adaptation of the conventional summary check column of the Doolittle solution.

TABLE II

Generalized Doolittle Presentation

Row	Operation	β_1	β_2	β_3	r_{1k}	r_{2k}	r_{3k}	S
(1)		1.0000	.3300	.2100	−1.0000			.5400
(2)		.3300	1.0000	−.4800		−1.0000		−.1500
(3)		.2100	−.4800	1.0000			−1.0000	−.2700
(4)	Repeat (1)	1.0000	.3300	.2100	−1.0000			.5400
(5)	Negative of (4)	−1.0000	−.3300	−.2100	1.0000			−.5400
(6)	Repeat (2)		1.0000	−.4800		−1.0000		−.1500
(7)	−.3300 times (4)		−.1089	−.0693	.3300			−.1782
(8)	(6) + (7)		.8911	−.5493	.3300	−1.0000		−.3282
(9)	−(8) divided by .8911		−1.0000	.6164	−.3703	1.1222		.3683
(10)	Repeat (3)			1.0000			−1.0000	−.27
(11)	−.2100 times (4)			−.0441	.2100			−.1134
(12)	.6164 times (8)			−.3386	.2034	−.6164		−.2023
(13)	(10) + (11) + (12)			.6173	.4134	−.6164	−1.0000	−.5857
(14)	−(13) divided by .6173			−1.0000	−.6697	.9985	1.6200	.9488
(15)	.6164 times (14)			−.6164	−.4128	.6155	*.9986*	.5848
(16)	(9) + (15)		−1.0000		−.7831	1.7377	*.9986*	.9531
(17)	−.2100 times (14)			.2100	.1406	*−.2097*	*−.3402*	−.1992
(18)	−.3000 times (16)		.3300		.2584	*−.5734*	*−.3296*	−.3145
(19)	(5) + (17) + (18)	−1.0000			1.3990	*−.7831*	*−.6697*	−1.0537

The general solution is read from rows (19) (16) (14) and is

$$\beta_1 = 1.3990\, r_{1k} - .7831\, r_{2k} - .6697\, r_{3k}\,.$$

$$\beta_2 = -.7831\, r_{1k} + 1.7377\, r_{2k} + .9986\, r_{3k}\,.$$

$$\beta_3 = -.6697\, r_{1k} + .9985\, r_{2k} + 1.6200\, r_{3k}\,.$$

which agrees, aside from the last place, with the result of the solution by determinants.

It is wise to check in the original equations (1), (2), (3) as soon as any β_i is found. Row (14), for example, should be checked by showing

$$(-.6697)(1.00) + (.9985)(\ .33) + (1.6200)(\ .21) = \ .0000$$

$$(-.6697)(\ .33) + (.9985)(\ 1.00) + (1.6200)(-.48) = -.0001$$

$$(-.6697)(\ .21) + (.9985)(-.48) + (1.6200)(\ 1.00) = \ 1.0001$$

The same should be done with row (16) as soon as it is computed. Row (19) should be treated similarly.

6. **Many regression equations.** If large numbers of regression equations are to be generated (the Strong Vocational Interest Study had 29 dependent variables), the following technique is suggested. Make a table with columns r_{1k}, r_{2k}, etc. and use the rows to indicate the different values of k. On another slip of paper insert the general values $\beta_1, \beta_2, \beta_3, \cdots \beta_n$ in successive rows so that a folding of the paper will bring any general β expansion in conjunction with the r's of any test, k. The scheme is illustrated in Table III.

TABLE III

No.	Occupation	r_{1k}	r_{2k}	r_{3k}		β_{1k}	β_{2k}	β_{3k}		r
1	Teacher	1.00	.33	.21		1.00	.00	.00		1.00
2	Physicist	.33	1.00	−.48		.00	1.00	.00		1.00
3	Office Worker	.21	−.48	1.00		.00	.00	1.00		1.00
4	Doctor	.17	.79	−.52		−.03	.72	−.17		.81
5	Lawyer	−.02	.16	−.59		.24	−.30	−.78		.64
6	Engineer	.16	.78	−.02		−.37	1.21	.64		.93
	β_1	1.3990	−.7831	−.6697		↑ 1.0000				
	β_2	−.7831	1.7377	.9986			↑ 1.0000			
	β_3	−.6697	.9985	1.6200				↑ 1.0000		
10	Mathematician etc.	.46	.96	−.49		.19	.82	−.14		.97

Thus, for the occupation of Engineer,

$$\beta_1 = 1.3990\ (.16) + (-.7831)(.78) + (-.6697)(-.02) = -.37$$

$$\beta_2 = -.7831\ (.16) + (\ 1.7377)(.78) + (\ .9996)(-.02) = \ 1.21$$

$$\beta_3 = -.6697\ (.16) + (\ .9985)(.78) + (\ 1.6200)(-.02) = \ .64$$

The value of the multiple correlation coefficient is then computed from the formula

$$r_{k.123\cdots n} = \sqrt{\beta_{1k} r_{1k} + \beta_{2k} r_{2k} + \ldots + \beta_{nk} r_{nk}}$$

In the illustration above

$$r_{k.123} = \sqrt{(-.37)(.16) + (1.21)(.78) + (.64)(-.02)}$$
$$= .93$$

7. **Regression equations by deletion.** The method of getting related regression coefficients and correlation coefficients, described by Kurtz [3], is also applicable. Again, a problem involving more than three variables is needed to show the real value of the scheme but the technique may be illustrated in the three variable case. We wish to find, from the forward solution of Table II, the regression equation and the multiple correlation coefficient when the first two fundamental variables only are used. We delete all columns involving test 3 and complete the back solution as indicated in Table IV, which may be viewed as a substitute for the last ten rows of Table II.

TABLE IV
(See Table II)

Row	Operation	β_1	β_2	β_3	r_{1k}	r_{2k}
(20)	Repeat (9)		-1.0000		$-.3703$	1.1222
(21)	$-.3300$ times (20)		$.3300$		$.1222$	$-.3703$
(22)	(5) + (21)	-1.0000			1.1222	$-.3703$

The results are

$$\beta_1 = 1.1222\, r_{1k} - .3703\, r_{2k}\,.$$
$$\beta_2 = -.3703\, r_{1k} + 1.1222\, r_{2k}\,.$$

and these agree with the results of section 4.

8. **The simplified back solution.** In every case in which the β's have been given in terms of r's the matrix of the coefficients is symmetric (sections 3, 4, 5, 7). One wonders if this symmetry is generally true and if it holds for normal equations of Type I or Type II.

Determinants are much more useful in establishing general properties, such as the one under discussion, than they are in computing the values of regression coefficients in the case of a problem involving many variables. We return to the determinant notation of section 3.

In each of the three types, and hence in the general case $d_{ij} = d_{ji}$ so that D is a symmetric determinant, $D_{ij} = D_{ji}$ and $\frac{D_{ij}}{D} = \frac{D_{ji}}{D}$. Hence the matrix of the coefficients of the solution is symmetric.

This result may be used (1) to check the expanded results or (2) to eliminate some of the work of the back solution. The n coefficients must be recorded for β_n after which the column indicated by r_{nk} may be dropped. The first $n - 1$ coefficients must be computed for β_{n-1} after which the column indicated by $r_{n-1,k}$ may be dropped, etc. The italicized entries in Table II are the ones which are eliminated in this way. The remaining coefficients are sufficient to completely determine the symmetric matrix.

The summary right hand check column can not be readily used in the simplified back solution but it is hardly to be recommended anyway. Kurtz [3] argues against it on the ground that it is not necessary. The essential check is to see that each β solution satisfies all of the original equations.

9. **Conclusion.** This paper provides a technique for the computation of general regression equations and shows how the technique may be combined with the Doolittle method in providing a practical means of mass prediction.

UNIVERSITY OF MICHIGAN.

REFERENCES

[1] TOLLEY, H. R. AND EZEKIAL, MORDECAI. *The Doolittle method for solving multiple correlation equations versus the Kelley-Salisbury Iteration method.* Journal of American Statistical Association, 1927(22), pp. 497–500.

[2] KELLEY, T. L. AND MCNEMAR, Q. *Doolittle versus Kelley-Salisbury iteration methods for computing multiple regression coefficients.* Journal of American Statistical Association, 1929 (24), pp. 164–169.

[3] KURTZ, A. K. *The use of the Doolittle method in obtaining related multiple correlation coefficients.* Psychometrika, Vol. I, no. 1, March 1936, pp. 45–51.

For historical and bibliographical references the reader is advised to consult:

GRIFFIN, H. D. *On partial correlation versus partial regression for obtaining the multiple regression equations.* Jour. of Ed. Psy. 1939(22), pp. 35–44.

GRIFFIN, H. D. *Simplified Schema for multiple linear correlation.* Jour. of Ex. Ed. Vol. I, no. 1, March 1933, pp. 239–254.

CONSTITUTION

Article I

NAME AND PURPOSE

1. This organization shall be known as the Institute of Mathematical Statistics.

2. Its object shall be to promote the interests of mathematical statistics.

Article II

MEMBERSHIP

1. The membership of the Institute shall consist of Members, Fellows, Honorary Members, and Sustaining Members.

2. Fellows shall be the only voting members of the Institute.

Article III

OFFICERS, BOARD OF DIRECTORS, COMMITTEE ON MEMBERSHIP, AND COMMITTEE ON PUBLICATIONS

1. The Officers of the Institute shall be a President, two Vice-Presidents, and a Secretary-Treasurer, elected for a term of one year by a majority ballot at the annual meeting of the Institute. Voting may be in person or by mail.

(a) Exception. The first group of Officers shall be elected by a majority vote of the individuals present at the organization meeting, and shall serve until December 31, 1936.

2. The Board of Directors of the Institute shall consist of the Officers and the previous President.

3. The Institute shall have a Committee on Membership composed of three Fellows. At their first meeting subsequent to the adoption of this Constitution, the Board of Directors shall elect three members as Fellows to serve as the Committee on Membership, one member of the Committee for a term of one year, another for a term of two years, and another for a term of three years. Thereafter the Board of Directors shall elect from among the Fellows one member annually at their first meeting after their election for a term of three years. The president shall designate one of the Vice-Presidents as Chairman of this Committee.

4. The Institute shall have a Committee on Publications composed of three Members or Fellows elected by the Board of Directors. The President shall designate a Vice-President as Ex Officio Chairman of this Committee.

Article IV

MEETINGS

1. A meeting for the presentation and discussion of papers, for the election of Officers, and for the transaction of other business of the Institute shall be held annually at such time as the Board of Directors may designate. Additional meetings may be called from time to time by the Board of Directors and shall be called at any time by the President upon written request from ten Fellows. Notice of the time and place of meeting shall be given to the membership by the Secretary-Treasurer at least thirty days prior to the date set for the meeting. All meetings except executive sessions shall be open to the public. Only papers accepted by a Program Committee appointed by the President may be presented to the Institute.

2. The Board of Directors shall hold a meeting immediately after their election and again immediately before the expiration of their term. Other meetings of the Board may be held from time to time at the call of the President or any two members of the Board. Notice of each meeting of the Board, other than the two regular meetings, together with a statement of the business to be brought before the meeting, must be given to the members of the Board by the Secretary-Treasurer at least five days prior to the date set therefor. Should other business be passed upon, any member of the Board shall have the right to reopen the question at the next meeting.

3. The Committee on Membership shall hold a meeting immediately after the annual meeting of the Institute. Further meetings of the Committee may be held from time to time at the call of the Chairman or any member of the Committee provided notice of such call and the purpose of the meeting is given to the members of the Committee by the Secretary-Treasurer at least five days before the date set therefor. Should other business be passed upon, any member of the Committee shall have the right to reopen the question at the next meeting.

4. At a regularly convened meeting of the Board of Directors, three members shall constitute a quorum. At a regularly convened meeting of the Committee on Membership, two members shall constitute a quorum.

Article V

PUBLICATIONS

1. In the beginning, the "Annals of Mathematical Statistics" shall serve as the official journal for the Institute. Other publications may be originated by the Board of Directors as occasion arises.

Article VI

EXPULSION OR SUSPENSION

1. Except for non-payment of dues, no one shall be expelled or suspended except by action of the Board of Directors with not more than one negative vote.

ARTICLE VII

AMENDMENTS

1. This constitution may be amended by an affirmative two-thirds vote at any regularly convened meeting of the Institute provided notice of such proposed amendment shall have been sent to each Fellow by the Secretary-Treasurer at least thirty days before the date of the meeting at which the proposal is to be acted upon. Voting may be in person or by mail.

BY-LAWS

ARTICLE I

DUTIES OF THE OFFICERS, BOARD OF DIRECTORS, COMMITTEE ON MEMBERSHIP, AND COMMITTEE ON PUBLICATIONS

1. The President, or in his absence, one of the Vice-Presidents, or in the absence of the President and both Vice-Presidents, a Fellow selected by vote of the Fellows present, shall preside at the meetings of the Institute and of the Board of Directors. At meetings of the Institute, the presiding officer shall vote only in the case of a tie, but at meetings of the Board of Directors he may vote in all cases. At least three months before the date of the annual meeting, the President shall appoint a Nominating Committee of three members. It shall be the duty of the Nominating Committee to make nominations for Officers to be elected at the annual meeting and the Secretary-Treasurer shall notify all Fellows at least thirty days before the annual meeting. Additional nominations may be submitted in writing, if signed by at least ten Fellows of the Institute, up to the time of the meeting.

2. The Secretary-Treasurer shall keep a full and accurate record of the proceedings at the meetings of the Institute and of the Board of Directors, send out calls for said meetings and, with the approval of the President and the Board, carry on the correspondence of the Institute. Subject to the direction of the Board, he shall have charge of the archives and other tangible and intangible property of the Institute. He shall send out calls for annual dues and acknowledge receipt of same; pay all bills approved by the President for expenditures authorized by the Board or the Institute; keep a detailed account of all receipts and expenditures, prepare a financial statement at the end of each year and present an abstract of the same at the annual meeting of the Institute after it has been audited by a Member or Fellow of the Institute appointed by the President as Auditor. The Auditor shall report to the President.

3. The Board of Directors shall have charge of the funds and of the affairs of the Institute, with the exception of those affairs specifically assigned to the President or to the Committee on Membership. The Board shall have authority to fill all vacancies ad interim, occurring among the Officers, Board of Directors, or in any of the Committees. The Board may appoint such other

committees as may be required from time to time to carry on the affairs of the Institute.

4. The Committee on Membership shall prepare and make available through the Secretary-Treasurer an announcement indicating the qualifications requisite for the different grades of membership.

5. The Committee on Publications, under the general supervision of the Board of Directors, shall have charge of all matters connected with the publications of the Institute, and of all books, pamphlets, manuscripts and other literary or scientific material collected by the Institute. Once a year this Committee shall cause to be printed in the Official Journal the Constitution and By-Laws and a classified list of all the Members and Fellows of the Institute.

Article II

DUES

1. Members shall pay five dollars at the time of admission to membership and shall receive the full current volume of the Official Journal. Thereafter, Members shall pay five dollars annual dues. The annual dues of Fellows shall be five dollars. The annual dues of Sustaining Members shall be fifty dollars. Honorary Members shall be exempt from all dues.

2. Annual dues shall be payable on the first day of January of each year.

3. The annual dues of a Fellow or Member include a subscription to the Official Journal. The annual dues of a Sustaining Member include two subscriptions to the Official Journal.

4. It shall be the duty of the Secretary-Treasurer to notify by mail anyone whose dues may be six months in arrears, and to accompany such notice by a copy of this Article. If such person fail to pay such dues within three months from the date of mailing such notice, the Secretary-Treasurer shall report the delinquent one to the Board of Directors, by whom the person's name may be stricken from the rolls and all privileges of membership withdrawn. Such person may, however, be re-instated by the Board of Directors upon payment of the arrears of dues.

Article III

SALARIES

1. The Institute shall not pay a salary to any Officer, Director, or member of any committee.

Article IV

AMENDMENTS

1. These By-Laws may be amended in the same manner as the Constitution or by a majority vote at any regularly convened meeting of the Institute, if the proposed amendment has been previously approved by the Board of Directors.

DIRECTORY OF INSTITUTE OF MATHEMATICAL STATISTICS

ACERBONI, ARGENTINO V., Larroque 232 Banfield, Buenos Aires, Republic of Argentina, South America.

ALTER, DINSMORE, Director of Griffith Observatory, Los Angeles, California.

AROIAN, LEO, Department of Mathematics, Colorado State College, Fort Collins, Colorado.

BACON, H. M., Box 1144, Stanford, California.

BAILEY, A. L., United Fruit Company, One Federal Street, Boston, Mass.

BAKER, G. A., Experiment Station, College of Agriculture, University of California, Davis, California..

BARRAL-SOUTO, JOSÉ, Córdoba 1459, Buenos Aires, Republic of Argentina, South America.

BEALL, GEOFFREY, 416 Queen Street, Chatham, Ontario, Canada.

BENNETT, A. A., Department of Mathematics, Brown University, Providence, Rhode Island.

BERKSON, JOSEPH, The Mayo Clinic, Rochester, Minn.

BERNSTEIN, FELIX, Department of Biometry, New York University, New York, New York..

BROOKS, A. G., Western Electric Company, Hawthorne Station, Chicago, Illinois.

BURGESS, R. W., Western Electric Co., 195 Broadway, New York, New York.

BUSHEY, J. HOBART, Department of Mathematics, Hunter College, Park Avenue and 68th Street, New York, New York.

CAMP, B. H., Department of Mathematics, Wesleyan University, Middletown, Conn..

CAMP, C. C., Department of Mathematics, University of Nebraska, Lincoln, Nebraska.

CARVER, H. C., Department of Mathematics, The University of Michigan, Ann Arbor, Michigan..

COLEMAN, E. P., Box 405, State College, Mississippi.

CRAIG, A. T., Department of Mathematics, The University of Iowa, Iowa City, Iowa..

CRAIG, C. C., 3020 Angell Hall, The University of Michigan, Ann Arbor, Michigan..

CRATHORNE, A. R., Department of Mathematics, The University of Illinois, Urbana, Illinois..

CROWE, S. E., 137 University Drive, East Lansing, Michigan.

DEMING, W. E., Fertilizer Investigations, Friendship Post Office, Washington, D. C..

DJOU, I. REN, 336 East Washington Street, Ann Arbor, Michigan.

DODD, E. L., Department of Mathematics, The University of Texas, Austin, Texas..

DOOB, J. L., Department of Mathematics, The University of Illinois, Urbana, Illinois..

DWYER, PAUL S., 407 Camden Court, Ann Arbor, Michigan.

EDGETT, G. L., Queen's University, Kingston, Ontario, Canada.

EISENHART, CHURCHILL, Department of Mathematics, The University of Wisconsin, Madison, Wisconsin.

ELSTON, JAMES S., The Travelers Insurance Company, Hartford, Conn.

EVANS, H. P., North Hall, The University of Wisconsin, Madison, Wisconsin.

FEINLER, F. J., 821 East C Street, Grants Pass, Oregon.

FELDMAN, H. M., Soldan High School, St. Louis, Missouri.

FERTIG, JOHN W. State Hospital, Worcester, Massachusetts.

FISCHER, C. H., Department of Mathematics, Wayne University, Detroit, Michigan.

FISHER, IRVING, 460 Prospect Street, New Haven, Conn..

FORSYTH, C. H., Department of Mathematics, Dartmouth College, Hanover, New Hampshire.

Foster, Ronald M., 122 East Dudley Avenue, Westfield, New Jersey.
Frankel, Lester R., 215 West 92nd Street, New York, New York.
Fry, T. C., 463 West Street, New York, New York..
Gavett, G. I., Department of Mathematics, The University of Washington, Seattle, Washington.
Girshick, M. A., Bureau of Home Economics, Department of Agriculture, Washington, D. C.
Glover, J. W., 620 Oxford Road, Ann Arbor, Michigan..
Grove, C. C., 143 Melburn Avenue, Baldwin, L. I., New York.
Hart, W. L., Department of Mathematics, The University of Minnesota, Minneapolis, Minnesota.
Henderson, Robert, Crown Point, Essex County, New York..
Hendricks, Walter A., National Agricultural Research Center, Beltsville, Maryland.
Henry, M. H., 829 Allegan Ave., East Lansing, Michigan.
Hotelling, Harold, Fayerweather Hall, Columbia University, New York, New York..
Huntington, E. V., 48 Highland Street, Cambridge, Massachusetts..
Ingraham, M. H., North Hall, The University of Wisconsin, Madison, Wisconsin..
Jackson, Dunham, 119 Folwell Hall, The University of Minnesota, Minneapolis, Minnesota..
Janko, Jaroslav, Praha XII, Kolínská 12, Czechoslovakia.
Keeping, E. S., University of Alberta, Edmonton, Alberta, Canada.
Kelley, Truman L., Lawrence Hall, Cambridge, Massachusetts..
Kullback, Solomon, Care of Department Signal Officer, Fort Shafter, Honolulu, T. H..
Kurtz, Albert K., Life Ins. Sales Research Bureau, Hartford, Connecticut.
Larsen, Harold D., Department of Mathematics, University of New Mexico, Albuquerque, New Mexico.
Leavens, Dickson H., 301 Mining Exchange Building, Colorago Springs, Colorado.
Lotka, A. J., One Madison Avenue, New York, New York..
Madow, William G., John Jay Hall, Columbia University, New York, New York.
Malzberg, Benjamin, New York State Department of Mental Hygiene, Albany, New York.
Mauchly, J. W., Department of Physics, Ursinus College, Collegeville, Pennsylvania.
Molina, E. C., 463 West Street, New York, New York..
Mudgett, Bruce D., Department of Economics, University of Minnesota, Minneapolis, Minnesota.
MacLean, M. C., Chief of Census Analysis, Dominion Bureau of Statistics, Ottawa, Canada.
McEwen, G. F., Scripps Institution, La Jolla, California.
Norris, Nilan, Department of Economics, The University of Maryland, College Park, Maryland.
Olds, E. G., 953 La Clair Avenue, Regent Square, Pittsburgh, Pennsylvania.
Ollivier, Arthur, Box 405, State College, Mississippi.
O'Toole, A. L., The University Club of St. Paul, St. Paul, Minnesota..
Ore, Oystein, Department of Mathematics, Yale University, New Haven, Connecticut.
Parente, A. R., 126 Church Street, Hamden, Connecticut.
Pixley, H. H., Department of Mathematics, Wayne University, Detroit, Michigan.
Pollard, H. S., Department of Mathematics, Miami University, Oxford, Ohio.
Reed, L. J., 615 North Wolfe Street, Baltimore, Maryland.
Rider, P. R., Department of Mathematics, Washington University, St. Louis, Missouri..
Rietz, H. L., Department of Mathematics, The University of Iowa, Iowa City, Iowa..
Romig, H. G., 463 West Street, New York, New York.
Roos, Charles F., Care of Mercer-Allied Corporation, 420 Lexington Avenue, New York, New York..

RULON, P. J., 4 Emerson Hall, Harvard University, Cambridge, Massachusetts.
SCARBOROUGH, J. B., Post Office Box 332, Annapolis, Maryland.
SCHULTZ, HENRY, Social Science Building, The University of Chicago, Chicago, Illinois..
SHEWHART, W. A., 158 Lake Drive, Mountain Lakes, New Jersey.
SHOHAT, J., Department of Mathematics, The University of Pennsylvania, Philadelphia, Pennsylvania..
SPIEGELMAN, MORTIMER, 325 West 86th Street, New York, New York.
STOUFFER, S. A., Department of Sociology, University of Chicago, Chicago, Illinois.
SWANSON, A. G., General Motors Institute, Flint, Michigan.
THOMPSON, WILLIAM R., 27 Oakwood Street, Albany, New York.
TOOPS, HERBERT A., The Ohio State University, Columbus, Ohio.
TRELOAR, ALAN E., Department of Biometry, The University of Minnesota, Minneapolis, Minnesota.
VICKERY, C. W., 906 East Monroe Street, Austin, Texas.
VAN DYK, PHYLLIS, 417 Stirling Court, Madison, Wisconsin.
WALKER, HELEN M., Teachers College, Columbia University, New York, New York.
WEIDA, F. M., Department of Mathematics, The George Washington University, Washington, D. C..
WELKER, E. L., 160 Mathematics Building, University of Illinois, Urbana, Illinois.
WHITE, A. E., Department of Mathematics, Kansas State College, Manhattan, Kansas.
WILKS, S. S., Fine Hall, Princeton University, Princeton, New Jersey..
WILSON ELIZABETH W., One Waterman Street, Cambridge, Massachusetts.
WRIGHT, SEWALL, Department of Zoology, University of Chicago, Chicago, Illinois..
WYCKOFF, J. F., Department of Mathematics, Trinity College, Hartford, Connecticut.
YONEDA, KIYOTAKA, Kwansei Gakuin, near Kobe, Japan.
ZOCH, RICHMOND T., 515 Jackson Avenue, Riverdale, Maryland.

Date of Issue	Date of Issue	Date of Issue

GIPNLK—H-40 I.A.R.I.—29-4-5—15,000